

COMBINING INFORMATION FROM BIOLOGICAL EXPERIMENTS VIA KALMAN FILTER

by

B. R. Murty, Indian National Academy of Sciences, New Delhi
and

W. T. Federer, Departments of Biometrics and Statistical Sciences, Cornell University

ABSTRACT

The Kalman filter or Kalman filtering procedure was developed by R. E. Kalman in the 1960's. It is a statistical procedure allowing use of prior (state of nature) information and smoothing techniques for parameter estimation in a model. The use of a prior is Bayesian in nature in that it contains the elements of Bayesian inference but Kalman filter goes further in fitting the model. It is useful for prediction, forecasting, fitting trends, and time series analyses for either past or future events. It can be used for linear as well as non-linear models and for normally distributed effects as well as non-normal cases. It has been used in many areas of investigation such as, e.g., agriculture, beer fermentation, biology, biotechnology, chemistry, clinical studies, economics, fisheries, forestry, growth studies, and a variety of other studies. It has been used as a monitoring device in medicine, space exploration, and other areas.

The first part of this paper considers methods used for combining data, when to combine data, and how to combine data using Kalman filtering. The nature of agricultural experiments, meta-analysis for plant breeding studies, a description of the Kalman filter procedure with an example, clustering procedures, some benefits of using Kalman filter, advantages of using Kalman filter in agriculture and related fields, some conclusions that may be drawn, and new developments associated with the theory and application of Kalman filter are discussed.

Part II of the paper deals with a literature review of applications of Kalman filter in a variety of fields. In forestry, it has been used to study crown growth and canopy structure, development of disease in stands, sustainability of a management system, trend monitoring in a land/forest cover, tree growth, weather relations to tree growth, and estimation of production. In hydrology, Kalman filter has been used to estimate missing hydrological data, estimate aquifer parameters from pumping data, estimate ground water flow problems due to soil properties, estimate ground water discharge and its contaminants, forecast ground water levels in aquifers, estimate catchment rainfall and run-off, forecast floods, and estimate water quality in rivers and streams. In fisheries,

Kalman filter has been utilized in aquatic systems for production models of fisheries, for abundance, biological dynamics, population parameters related to harvest, for prediction of stockage, productivity of sub-populations, for role of environment on the population dynamics, estimation of size, productivity, and stock biomass of fish stocks, and for estimation of growth and mortality. The application of Kalman filtering in biotechnology for the past two decades is in the area of on-line estimation of process control, estimation and prediction of biomass, growth rates, product formation, and enzymatic action detection of important state variables in bio-processes and control of growth of cells. In biochemistry and analytical chemistry, Kalman filter is used for the simultaneous detection and estimation of groups and compounds, for dynamics of industrial processes, for estimation and differentiation of kinetics of isozymes and other compounds, for resolution of complex or overlap, for detection and correction in fluorescence response, and for analysis of ultra-violet and NMR spectra. Kalman filtering has been applied in various medical fields such as signal processing (ECG, EEG signals, etc.) in monitoring patient progress as in renal transplants, in analysis of complex data found in hemodynamics and vision problems, and in trend detection and analysis of spectra to name a few. In agriculture, Kalman filter has found limited use even though there is great potential in this area for its application. It has been used for the estimation and prediction of productivity of crops, forests, animals, poultry, and fisheries, for an analysis of soil moisture, nutrient, and productivity status, for monitoring the production and management of crops and animals, for studying the dynamics of disease and pest build-up and control, for monitoring process and quality of food products, for trend analysis of production, for studying the effect of climatic changes, for predicting supply/elasticity and demand, and for studying effects of smoothing in biological and econometric models. The numerous applications of Kalman filter in engineering and physics are not included in the present discussion.

COMBINING INFORMATION FROM BIOLOGICAL EXPERIMENTS VIA KALMAN FILTER

INTRODUCTION

Combining information from sets of trials over time and space on independent and related studies, not necessarily identical, is important in the biological, physical and engineering sciences. Such an analysis could consist of discovery of meaningful patterns by accounting for variability across experiments (Carlin, 1992; Gaver *et al.*, 1992). Summarization of the information using reliable estimates allows inferences to be drawn over a broader spectrum and generation of hypotheses for future studies. The power of such meta-analysis is best illustrated by the aspirin-heart attack data which detected the positive effect of aspirin not detected earlier (Canner, 1987; Gaver *et al.*, 1992; Higgins, 1996). Difficulties encountered are unbalanced data, changes in treatment composition and protocol from experiment to experiment, and changes in variables and subjects. These prevent the use of standard statistical procedures. The increased use of meta-analysis employing modified Bayesian approaches and state-space models is of recent origin. Matching strategies for combining data from different variables is being followed in medical, environmental, and molecular biology experiments as they are expensive to repeat or non-repeatable at all. Combining information helps to utilize optimally available data and cost efficiency strategies. In agriculture where multi-location trials are of common occurrence, strategies for regional/environmental specific knowledge, for clustering of environmental sites and treatments (McLachlan and Basford, 1988; Gauch, 1988), monitoring genetic advance over time in plant breeding, predictions of future performance by estimating trends in production, consumption and pricing in economics (Normand and Trichler, 1992, and progress of medical treatments (Mosteller and Chalmero, 1992), spatial analyses proposed by Federer (1998) and Federer *et al.* (1998), space sciences (Leondes, 1992), and also in combining information with state-space modeling is possible.

METHODS USED

Among the methods commonly used for combining data are signal processing methods using dynamic linear or nonlinear models (Brown, 1983), variance component analysis of combined data using fixed, random, and mixed effect models (Federer, 1951; Sprague and Federer, 1951; Yates and Cochran, 1938), pooling across studies with weighting of variances (Verducci, 1988; Gregoire and Walters, 1989), record linkage and merging of files, P-value based methods (Fisher, 1932), general Bayesian estimation in time series (Givens, Smith, and Tweedie, 1997), matching information from multiple sources as for DNA, and use of state-space models such as Kalman filter, which is considered to be one of the significant discoveries of the 20th century (Grewal and

Andrews, 1993). Among these methods, Kalman filter is a special case of time-series models called dynamic linear models and is a recursive procedure for combining prior observations with current observations in an updating algorithm. In the analysis of trends, forecasting seasonal adjustment in trends, corrections, and smoothing in abrupt changes of trend in time series, various methods of Kalman filter are popular and efficient. Epidemiology of pathogens and pests over time in crops, animals, and humans can also be analyzed using Kalman filter and related methods. Perennial crop production, animal production, and dynamics of aquatic organisms as fish can be predicted using Kalman filter and any missing data can be estimated, and smoothing procedures can be adopted for examining the trends. A comprehensive review of this subject is made by the National Research Council (NRC) Panel on statistical issues and opportunities in research in the combining of information (Gaver, Draper, Croll, Hedges, Morris, and Waternaut, 1992 in the NRC publication, Grewal and Andrews, 1993, Olkin and Sampson, 1998, van den Bergh, Bulton, Nijkamp, and Peppring, 1997). A detailed analysis of why to combine information, how to combine data, nature of the agriculture experiments, methods of analysis comparing state space models with other methods, application of Kalman filter with examples in several fields, and recent developments in the theory and application of Kalman filter will be presented in subsequent pages.

PART I - COMBINING INFORMATION

Why Combine Information?

- * Pooling data of different conditions and precisions to yield more accurate conclusions (random sub-populations)
- * Drawing more general conclusions
- * Establishing reference values
- * Analysis of trends over time (examples: disease, demography, economics, environment, failure rates, product degradation)
- * Forecasting (control variable effectiveness)

When to Combine (Criteria)

The following criteria are to be satisfied when combining experimental information is being considered:

- * Information sources similar enough for realism (repeated surveys, production, money supply, power system load adjustments, seasonal adjustments/price elasticity, opinion polls)
- * Ability to express similarity within the model framework ((linking data from several sources)
- * Careful examination of outliers (outliers versus changes in levels (signals) and slopes (trends) and also to explain abrupt changes in slope)
- * Traditional and rank analyses applicable
- * Combining with multivariate techniques - reduction of dimensionality

- * Gibbs sampling - linear and nonlinear time series
- * Combining information from different sources, each with its own statistical properties (sub-populations with different means, variances and/or other parameters)

How to Combine Information

- * Pooling across studies: Simplest form of combining data.
 - homogeneous studies

$$d_i = \delta + e_i \quad \hat{d}_w = \sum_{i=1}^k w_i d_i / \sum_{i=1}^k w_i, \quad w_i = 1/v_i,$$

where \hat{d}_w is a pooled estimate of \hat{d} and v_i is the within study variance.

- heterogeneous studies (borrowing strength across studies)

differences in d_i	differences between studies
(within study variance)	(variation in treatments or protocol)

$$d_i = \delta_i + e_i,$$

$$\delta_i = \delta + \epsilon_i,$$

and

$$\delta_i = (1 - \hat{\beta}_i) d_i + \hat{\beta}_i \delta_w, \quad \beta = v_i / (v_i + \hat{\tau}_i^2).$$

$$\hat{\tau}_i^2 = \{ \sum w_i^2 (y_i - \hat{\mu})^2 - v_i \} / \sum w_i^2 \text{ for } w_i^2 = 1 / (v_i + \hat{\tau}_i^2)$$

is the between studies variance and can be found by iteration. w_i is the reciprocal of the sum of the within and between studies variance. This is the simplest form of combining data.

- * Combining across variables (sites, zones, sources, agricultural experiments)
 - prediction/classification /dimensionality reduction (principal component analysis, PCA) may also be done depending on the model
- * Combining across subjects (varieties)
 - age groups, ethnic groups, social strata
 - is a common statistical link between social and agricultural studies
- * Combining across time
 - is done in production, trend, and agricultural experiments over time
- * Combining data sets from different studies using Bayesian approach and Kalman filter
 - is done in environmental, medical, forestry, fisheries, and is possible in agricultural studies
- * Techniques for robustness through sensitivity analysis after combining
 - relationship between conclusions and models (Iyengar and Greenhouse, 1988)
- * Matching strategies as in DNA profiling and electrophoresis studies in biochemistry and environmental studies

- * Estimating missing observations
 - important component of combined analyses where filtering methods like Kalman filter are used
 - intervention analysis is based on possible abrupt changes in trend
 - efficient forward and backward recursive methods needed and for search of possible anomalies in a small section of the data with outliers
 - implemented via Kalman filter and demonstrated in several situations

NATURE OF AGRICULTURAL EXPERIMENTS

Methods for combining information from several sources (meta-analysis) for agricultural studies is needed for the following reasons:

- * Unbalanced treatment composition within and across seasons and different local checks
 - genotype composition changes over years
- * Changes in sites
- * Change in spectrum of genotypes for specific mega-environments
- * Quality of data is variable
- * Data on limited variables such as height, maturity, cob weight, disease, yield, etc. but incomplete data on some variables only available in some locations
- * Grouping of locations using prior information on soils, rainfall, cropping system
- * Multi-location trials over years via multivariate analyses

META-ANALYSIS FOR PLANT BREEDING STUDIES

Combining information across several sites by meta-analysis procedures allows for:

- * Reduction in the number of trial sites and use of ancillary data from local trials and demonstration studies
- * Measuring rate of advance over checks
- * Analysis for adaptation over regions
- * Overemphasis on one variable such as yield reduced by including more variables
- * Need for adequate appreciation of use of prior information
 - problems are analogous to medical, environmental, and space studies

The objectives for plant breeders are

- * Continuous genetic enhancement of material
 - rate of progress in yield, diseases, pest build-up, change of system
- * Development of region specific technology
- * Rationalization of sites
 - grouping sites and identification of key sites using clustering methods and D^2
- * Genetic enhancement for specific adaptation

- variance-covariance structure comparisons, homogeneity, PCA, D²
- small samples, high dimensionality

* Understanding mechanisms of adaptation

Meta-analysis is a statistical procedure for achieving these objectives.

METHODS OF COMBINING INFORMATION

One method is Bayesian analysis. A general formulation of a Bayesian procedure follows.

* It is assumed that the parameter τ , the between site or study standard deviation, is known. Then, the following steps are followed:

1. An initial joint normal prior density for μ , the population mean, and σ , the population standard deviation, is assumed.
2. The data of past sets of observations are used to obtain an initial posterior density for μ and σ .
3. The current data set is used to obtain a posterior density α_{I+1} and a revised density for μ and σ .

* The revised posterior densities are used to obtain point and interval estimates for $Y_{I+1,K+1}$ for various assumptions about the knowledge of μ , σ , and τ .

An example is used to to illustrate the approach. Data for this example are presented in Table 1. The model is written as

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

$i = 1, \dots, I$, $I+1 = 4$ locations for $I = 3$, $j = 1, \dots, J$ for $J = 10$ varieties, $K = 9$ in $I+1$, α_i are distributed as $N(0, \tau^2 \sigma^2)$ and ϵ_{ij} are distributed as $N(0, \sigma^2)$. τ is set equal to one as an initial value to start the iteration.

$$a = K\tau^2/(1 + K\tau^2) = 9(1^2)/(1 + 9(1^2)) = 0.90$$

$$b = IJ\tau^2/(1 + J\tau^2) = 3(10)(1^2)/(1 + 10(1^2)) = 30/11 = 2.73$$

$$\hat{\mu} = (a\bar{y}_s + b\bar{y}_c)/(a + b) = \{0.90(87.67) + 2.73(90.90)\} = 90.09$$

where $\bar{y}_s = 87.67$ and $\bar{y}_c = 90.90$ for s for same site and c for Cousin data from other sites, respectively.

$$\alpha_4 = a(\bar{y}_s - \hat{\mu}) = 0.90(87.67 - 90.09) = -2.18$$

The predicted value for V10 at location 4 is

$$Y_{4,10} = \hat{\mu} + \alpha_4 = 90.09 + (-2.18) = 87.8$$

The estimated variance is obtained as

$$\sigma^2 = \{SSE_c + SSB_c/(1 + J\tau^2) + \sum_{j=1}^k (Y_{i+1,j} - \bar{y}_s)^2 + ab\tau^2(\bar{y}_s - \bar{y}_c)^2/\nu(a + b)\} = 4.91,$$

where σ^2 is distributed as χ_ν^2 where $\nu = IJ + K - 2 = 3(10) + 9 - 2 = 37$.

$$\sigma = \sqrt{4.91} = 2.43.$$

Treatment	Location			
	1	2	3	4
V1	88.0	85.9	94.2	87.5
V2	88.0	88.6	91.5	85.5
V3	94.8	88.5		
V4				
V5				
V6				
V7				
V8				
V9	89.0			86.2
V10	93.0			?
Mean	90.4	88.4	93.8	87.2

Table 1. Sketch of data for $K = 10$ entries at $I + 1 = 4$ sites. Data for $Y_{4,10}$ to be estimated.

KALMAN FILTER

The Kalman filter

- * is a recursive procedure for combining prior observations with current observations via an updating algorithm,
- * estimates unknown parameters in a linear dynamic model which is specified by two sets of equations, i.e., a measurement model (each site) and a system model (over sites),
- * involves least squares method and stochastic differential equations,
- * allows optimal prediction of future states ($t = 1, 2, \dots, t$), and
- * permits model mixing for versatility, and
- * is a method of updating by using Kalman gain which is the optimal gain derived by minimizing the error of $(\theta - \hat{\theta})$.

Kalman filter may be used for multi-location varietal trials to

- * predict local check performance at other sites,

- * compare relative performance of local checks (based on the pooled observations of several local checks),
- * characterize sites and among site comparisons in a better way,
- * account for changing varietal composition over time (years), and
- * estimate trends in gain over groups of sites and/or time.

The basic mathematics underlying the Kalman filter in matrix form is succinctly explained by Meinhold and Singpurwalla (1983) and is given below. Let the t^{th} observation be

$$Y_t = F_t \theta_t + v_t \quad (\text{data}) \quad (\text{measurement equation}) \quad (1)$$

where θ_t , a $IJ + K$ column vector of parameters, is analogous to β in linear regression at time $t = 1, 2, \dots, IJ + K$ but θ_t changes with time t . Let

$$\theta_t = G_t \theta_{t-1} + w_t \quad (\text{state equation or systems equation}) \quad (2)$$

The state variable θ_t completely describes the behavior of the dynamics of the system. θ_t is subject to white noise w_t . $\hat{\theta}_t$ is an estimate of θ_t in terms of Y which is measured as given in equation (1). The distribution of v_t and w_t is independent as explained in the following. From Bayes theorem we know that

$$\begin{array}{ll} \text{Prob}\{\text{state of nature}|\text{data}\} & \propto \text{Prob}\{\text{data}|\text{state of nature}\} \\ \text{Prob}\{\theta_t | Y_t\} & \propto \text{Prob}\{Y_t | \theta_t, Y_{t-1}\} \times \text{Prob}\{\theta_t | Y_{t-1}\} \\ (\text{posterior distribution at } t) & (\text{Likelihood}) \quad (\text{prior distribution of } \theta_t) \end{array} \quad (3)$$

It is assumed that $(\theta_t | Y_t)$ is normally distributed with mean $\hat{\theta}_{t-1}$ and covariance Σ_{t-1} . Then, $(\theta_{t-1} | Y_{t-1})$ is normally distributed with mean $G_t \hat{\theta}_{t-1}$ and with covariance $R_t = G_t(\Sigma_{t-1})G_t' + w_t$ using equation (2) because $\theta_t = G_t \theta_{t-1} + w_t$. If X , say, is normally distributed with mean μ and covariance Σ , this implies that CX is normally distributed with mean $C\mu$ and covariance $C\Sigma C'$. Now $\hat{e}_t = Y_t - \hat{Y}_t = Y_t - F_t G_t \theta_{t-1}$, and F_t , G_t , and θ_{t-1} are known. Then

$$\text{Prob}(\theta_t | Y_{t-1}) = \text{Prob}(\theta_t | e_t, Y_{t-1}) \propto \underset{\text{likelihood}}{\text{Prob}(e_t | \theta_t, Y_{t-1})} \text{Prob}(\theta_t | Y_{t-1})$$

where θ_t is an unknown vector of parameters, $t = 1, 2, \dots, IJ + K$, F_t is a known row vector, $\theta_t' = [\mu, \beta_t]$, $\beta_t = \alpha_j$ if Y_t is from j , and $F_t = [1, 1]$; v_t is a random observation vector $N(0, V_t)$, w_t is a random variable $N(0, \tau^2 \sigma^2)$, G_t and K_t are known matrices, w_t is $N(0, W_t)$, and v_t and w_t are independent.

The advantages of Kalman filter over the empirical Bayesian approach described above is that although it is related to a Bayesian procedure, it is applicable to both time and non-time series and considers the posterior distribution (observation equations) along with a systems approach (system equations). Conceptually, Kalman filter is based on a Bayesian approach as was illustrated by Meinhold and Singpurwalla (1983). Kalman filter is an optimal procedure for obtaining the estimate of θ by minimizing the error $(\hat{\theta} -$

θ) in the least squares sense with $(\hat{\theta} - \theta) \rightarrow 0$ as $t \rightarrow \infty$. Thus, one way is to view statistically the Kalman filter as a problem in Bayesian inference and updating procedure that consists of a prior guess about the state of nature and then adding a correction to this guess, the correction being determined by how the guess performed in predicting the next observation. As explained by Meinhold and Singpurwalla (1983), Kalman filter can be viewed as the evolution of a series of regression functions of G_t on e_t at times $0, 1, \dots, t-1, t$ each having a different intercept and regression coefficient and the evolution stems from a learning process involving all of the data.

To illustrate the use of Kalman filter to combine a series of experiments in space and/or time, we proceed as follows. From equations (1) and (2), we set

$$G_t = \begin{pmatrix} 1 & 0 \\ 0 & \delta \end{pmatrix} \quad K_t = \begin{pmatrix} 0 & 0 \\ 0 & 1-\delta \end{pmatrix} w_t = \begin{pmatrix} 0 \\ w \end{pmatrix} \quad \text{and} \quad W_t = \begin{pmatrix} 0 & 0 \\ 0 & \tau^2 \sigma^2 \end{pmatrix},$$

where w is a random variable $N(0, \tau^2 \sigma^2)$, $\delta = 1$ if t and $t-1$ observations are from the same location (site) and 0 otherwise, and τ and σ are given. Then

$$\theta_t = G_t \theta_{t-1} + K_t w_t$$

Now iterate from $t = 1$ to $t = IJ + K$ with the following equations (4) to (7). Let $Y_t = (y_1, y_2, y_3, \dots, y_t)$; Y_t , W_t , and $V_t = (v_1, v_2, v_3, \dots, v_t)$ are known. The posterior distribution of θ_t is $N(\theta_t, \Sigma_t)$ (Meinhold and Singpurwalla, 1983). Then,

$$\hat{\theta}_t = E[\theta_t | Y_t, V_t, W_t] = G_t \hat{\theta}_{t-1} + R_t F_t' (V_t + F_t R_t F_t')^{-1} \hat{e}_t \quad (4)$$

and

$$\Sigma_t = V(\theta_t | Y_t, V_t, W_t) = R_t - R_t F_t' (V_t + F_t R_t F_t')^{-1} F_t R_t \quad (5)$$

given that

$$R_t = G_t \Sigma^{-1} G_t' + K_t W_t K_t' \quad (6)$$

and

$$\hat{e}_t = Y_t F_t G_t \hat{\theta}_{t-1} \quad (7)$$

The procedure is started by selecting the initial values θ_0 and Σ_0 , and iterating equations (4) through (7) up to the last data point. An interval estimate of the $I+1, K+1^{\text{th}}$ observation is obtained as

$$Y_{I+1, K+1} = \mu + \beta_{IJ+K} \pm Z_{(1-\nu)/2} \sqrt{1 + p_0} \quad (8)$$

where p_0 is the sum of all elements of $IJ+K$; this is a $100\nu\%$ confidence interval. To obtain an estimate of a future observation, use $\mu + \beta_{IJ+K}$. Thus, a statistical way to view Kalman filter is to consider it as an updating procedure that consists of a prior guess about the state of nature and then to add a correction to this guess. The correction is determined by the performance of the guess in predicting the next observation. As

explained by Meinhold and Sinpurwalla (1983), Kalman filter can be viewed as an evaluation of a series of regression functions of θ_t on e_t at times $t = 0, 1, 2, \dots, t-1$, each having a different intercept and regression coefficient. The evolution stems from a learning process involving all the data.

Using the data of Table 1, a Kalman filter estimate was obtained and is compared with the Bayes estimate obtained previously. The results are:

	$Y_{4,10}$	σ
Kalman filter	86.7	1.92
Bayes	87.8	2.43

The Kalman filter estimate has greater precision by $1 - 1.92/2.43 = 21\%$. If μ and σ are unknown, the problems become more complex. A computer program for obtaining Kalman filter analyses is available from the Institute of Mathematical Statistics, Washington, D. C.; also given in Math Lab which is available to all universities.

CLUSTERING PROCEDURES

Using a state-space model to cluster-like observations together allows us to

- * use Kalman filtering to estimate values of local checks in other sites for such characteristics as height, maturity, cob weight, disease complex, and yield (This produces a matrix of values of the characteristics.),
- * use last year's estimate of a response as the initial estimate of $\hat{\theta}_t$ separately for checks and for all other entries, and for the best entry of last year,
- * utilize changes in θ_t and its variance to group into clusters using the data on non-check entries only or perhaps on all entries,
- * use principal component analysis to compare covariance matrices of the clusters so formed,
- * compare the new clusters over these based on empirical geographical grouping such as soil, rainfall, fertility, and clay mineralogy data,
- * use generalized distance analysis (D^2 -statistic) of the locations using all characteristics with or without yield and compare clusters formed with those obtained as above,
- * estimate θ_t for maturity, disease resistance, and other characters which are of interest, and
- * use to detect abrupt changes such as outliers and for smoothing.

Some of the problems encountered when using any one of the many clustering procedures are:

- (i). Most clustering procedures are descriptive and the results depend heavily on the vagaries of the methods used (Gnanadesikan *et al.*, 1984). This is a source of confusion for the user.
- (ii). Multivariate classification approach is better than others but multivariate normality and constancy of covariance matrices are assumed.

- (iii). Methods do not allow for noise or data points that do not fit into the prevailing cluster pattern.
- (iv). A state-space model will account for the problem in (iii) and to some extent to the one in (ii).
- (v). There are possible negative biases in estimates of variances in meta-analysis due to sample size differences and variance heterogeneity (Li, Shi, and Roth, 1994).

Using Tocher's method with the D^2 statistic, tests of significance for intra-cluster and inter-cluster D^2 s are possible.

Multivariate techniques may be used to cluster sites and the results compared with those obtained from Kalman filtering. A principal component analysis could be used on variables such as yield, maturity, height, lodging, disease, etc. to obtain an empirical grouping via PC1, PC2, and perhaps PC3. The variance-covariance matrices of genotype by environment interaction from clusters of sites may be compared. A measure of stability of groupings over years may be obtained to help explain the resulting biological differences. Crucial variables for discrimination between groups may be identified. Principal component analysis can be used to characterize clusters of sites and any ancillary information accumulated can be used for further discrimination among clusters.

Several illustrations of the efficiency and use of Kalman filter are presented in Tables 2, 3, 4, 5 and 6. Each of these situations are discussed below. The data in Table 2 represents a diversity of reported studies comparing Kalman filter with traditional methods of estimation. Tables 3 and 4 represent an analysis of a plant breeding yield trial of maize over 20 locations and clustering done by Kalman filtering approach. Table 5 represents meta-analysis of trends in acute myocardial infection data comparing three different methods of trend analysis. Table 6 illustrates the use of Kalman filter in linking data across systems.

An application of Kalman filter approach was used on 20 locations to group the 20 sites into clusters (groups) using all India maize trial data. The location of the 20 trials was four in the Indo-Gangetic Belt (Delhi and Punjab), five in the Uttar Pradesh area, three in the Bengal area, five in the Rajasthan and Madhyr Pradyh area, and three in the Deccan area. There were 18 entries (genotypes) in each trial which were made up of two standard checks, one local check, and 15 advanced entries. Six groups of locations resulted from the Kalman filter analysis described below with group sizes of four (Tribal), three (Diara), two (Hill), four (Western U P), four (P, H, D), and three (Deccan). The estimates of θ_t and the standard errors for checks, advanced entries, and the best entry over clusters are presented in Table 3. The mean yield of checks varied from 1059 to 1491 kg/ha, location means varied from 1999 to 4113 kg/ha, averages of advanced entries varied from 2035 to 4835, and the means of the best entry varied from 2057 to 5723 kg/ha. The latitudes of the locations ranged from 11° north to 29° north, representing a diversity of ecological regions.

The yield data from locations was arranged in a sequence of locations with increasing location means. Kalman filter procedure was applied for this sequence with estimation of θ_t , which is the rate of yield response of genotypes over locations. Starting iteration and initial estimate of θ , based upon previous year's data using the regression coefficient of yield of genotype on location means as in Finlay and Wilkinson (1963) as

Subject	Traditional		Kalman filter	Remarks
	estimate	estimate		
* (1) Nation Opinion Poll				
% conservative	47.3 (2.39)	47.9 (1.68)		30.% red.. in variance
Repeated surveys	49.7 (5.88)	47.8 (2.96)		49.6% red. in variance
+ (2) Financial time series	0.81 (0.15)	0.84 (0.12)		0.90 = true parameter
(Weekly data)	0.92 (0.09)	0.93 (0.07)		0.95 = true parameter
	0.12 (0.08)	0.11 (0.04)		0.10 = true parameter
x(3) Price change (cents/gal.)	$Y_{157} = -10.40$	-8.36 (2.86)		Superior to LSE
in gasoline (1978-91)	$Y_{158} = -6.10$	-2.51 (0.59)		Superior to LSE
t=1, 156 used				
* (4) Relation of relative				
income (z1), price of				Kalman filter superior by 28%
spirit (z2) to consumption				influence of i^{th} observation, fixed interval smoothing
Simple regression (1870-1938)				model validation done
(5) Labor force surveys				
Number hours/week				Kalman filter variance = 0.15
				Sampling variance = 0.39
(6) Milk production USA (1970-82)				Backward shift estimation for 1976 and 1979
* Model validation also done				
+ Corrected for correlation between measurement and state errors				
x Bilinear time series - Gibbs sampling used				
References: (1) Abraham and Vijayan (1992), (2) Papanastassiou (1992), (3) Chen				
(1992), (4) Kohn and Ansley (1987), (5) Pfeiffermann (1991), (6) Normand and Trichler				
(1992)				
LSE = least squares estimate				

Table 2. Some examples of the efficiency of Kalman filter over other procedures. Values in parentheses are standard errors of the estimates.

an initial value of θ , iteration for θ_t was done for the sequence of locations of the present data. When the estimate of θ_t increased by 1.65σ of the estimate, the next cluster is started. The estimates of θ_t and their variance for checks, advanced entries, and the best advanced entry over the clusters and the six clusters so formed are given in Table 3, along with the initial estimates of the θ 's used for the iteration. As provided in the Kalman filter procedure, past observations over locations in each cluster was combined in the set observations of the next location; thus $\hat{\theta}_t$ is an estimate of the cumulative data.

	Cluster						
	Initial*I		II	III	IV	V	VI
Checks - θ_t	0.28	0.30	0.34	0.38	0.39	0.39	0.40
- standard error	0.107	0.010	0.015	0.017	0.018	0.018	0.019
Advanced entries - θ_t	0.85	0.89	0.95	1.01	1.07	1.16	1.22
- stand. err.	0.095	0.018	0.021	0.023	0.025	0.030	0.034
Best entry - θ_t	0.86	0.97	1.02	1.08	1.17	1.20	1.24
- standard error	0.087	0.016	0.020	0.021	0.023	0.028	0.031
Number of locations/cluster	4	3	2	4	4	3	
	Tribal Diara Hill UPplain nw India Deccan						

* Initial value of θ_t was based on previous year's data.

Table 3. Rates of yield response (θ_t) and its standard error of genotypes over locations and clustering of locations (I to VI) via Kalman filter.

The six clusters formed by Kalman filter could be rationally explained ecologically. Tribal areas are with least inputs. Diara is area cultivated after floods receded. Hill is the Himalayan region. U. P. plains is the Gangetic delta. N. W. India comprising Punjab, Delhi, and Haryana represented Indo-Gangetic Northwest Plains. Deccan is the Deccan peninsula. The differences between the clusters were examined by principal component analysis of the data on five variables, i.e., days to silk, days to maturity, disease score, height of ear, and yield from the same trial and are presented in Table 4.

Principal component	Cluster						Mean
	I	II	III	IV	V	VI	
PC1 %	62	50	70	68	75	70	60
PC2 %	20	18	21	17	16	12	19
PC1 + PC2 %	82	68	91	85	91	82	79
PC3 %	7	11	5	8	5	9	10
PC1 + PC2 + PC3 %	89	79	96	93	96	91	89

Table 4. Principal component analysis of clusters for five variables. Percent of variation accounted for by each of three components.

From the results in Table 4, we note that a principal component analysis is useful in understanding cluster differences. The first two principal components accounted for a large proportion of the total variance which suggests relative homogeneity within each cluster except for cluster II. This demonstrates the utility of Kalman filter in plant

breeding. Cluster II variance was not accounted by PC1 + PC2 + PC3 as much as in the other clusters. Intra-cluster diversity can be explored further. For this, the variance-covariance matrix, Σ , for each cluster needs to be examined as there may be variance-covariance heterogeneity present. The effect of such heterogeneity on cluster orientation, size and shape needs to be examined. A reparameterization of Σ in terms of the eigenvalue decomposition is recommended, i.e., $\Sigma = D_k \Delta_k D_k'$. There is also a need to test vector means to explain cluster variation.

An example of combining information for trend analysis in clinical studies is presented by Rao and Shimitzu (1992), who give an analysis for data obtained in the National Health Discharge Survey in 1986 and combined over ten years. They performed a meta-analysis for trend estimation and compared three different methods of trend analyses, i.e., linear trend analysis, variance component estimation, and a Bayesian analysis. The variable is acute myocardial infection (AMI). There were three different hospitals representing regions for which sample discharge information was obtained. Let a_t be the number of cases of AMI, n_t is the sample number of discharges out of N_t total discharges, $\hat{P}_t = a_t / n_t$, $V(\hat{P}_t) = (1 - f_t)[\hat{P}_t (1 - \hat{P}_t)] / (n_t - 1)$, and $f_t =$ not obtained. The linear trend method, A in Table 5, utilizes the following:

$$Y_{ij} = \mu_t + \epsilon_{ij} ,$$

$$\bar{y}_t = \alpha + \beta a_t + \bar{\epsilon}_t ,$$

and

$$\bar{y}_c = \alpha + \beta x_t + \bar{\epsilon}_t .$$

A weighted least squares estimate of α and β was obtained. The second trend method, B in Table 5, used was variance components using MINIQUE solutions for the variance components. The model was a one-way classification using the linear model

$$Y_{ij} = \mu + \alpha_t + \epsilon_{ij}$$

and obtaining solutions for the among and within hospital variance components. The third method, denoted as C in Table 5, is an empirical Bayes procedure. Y_{ij} is $N(\mu, \sigma_t^2)$ and the prior distribution on μ_t is assumed to be $N(\mu, \sigma_t^2)$. Then the Bayes estimate of μ_t is

$$B_t = (1 - \hat{a}_t)\bar{y}_t + \hat{a}_t \hat{\mu}_t$$

A comparison of the three methods is given in Table 5. A Kalman filter estimate would more than likely be more precise than any of these methods as the Bayesian approach was better than the other two.

Kalman filter in linking data across systems

Kalman filter is useful in systems analysis for linking data across systems as, for example in the following cases:

Hospital	\hat{P}_t	Trend method		
		A	B	C
1 - mean	0.015	0.020	0.0224	0.0224
- confid. int.	± 0.007	± 0.005	± 0.003	± 0.0003
2 - mean	0.011	0.0162	0.0204	0.0203
- confid. int.	± 0.005	± 0.0036	± 0.0031	± 0.0003
3 - mean	0.0321	0.0319	0.0304	0.0309
- confid. int.	± 0.0060	± 0.0038	± 0.0028	± 0.0037

\hat{P}_t is the sample proportion, A is the linear trend estimate, B is the variance component estimate, and C is the empirical Bayes estimate.

Table 5. Meta-analysis of trend estimation in acute myocardial infarction by three different methods, Rao and Shimitzu (1992).

- * Data on systems with different structures but some parts in common such as, for example, radio and air conditioner, train and automobile, computer and satellite, etc.
- * Failure, detection, and identification of the component
- * Collection of data on failure of a component and failure time
- * Estimation of individual component failure rates
- * Estimation of system reliability and its use for other systems

For example, an electromagnetic valve has an AC component and an AC activator. The failure rate of the two components and their 95% confidence intervals are given in Table 6. The confidence intervals for the failure rates of the two components are much smaller for the Kalman filter than for the least squares estimate.

<u>Least squares estimate</u>	<u>Kalman filter estimate</u>
$\lambda_1 = 0.057 \text{ } 0.031$	$\lambda_1 = 0.051 \text{ } 0.020$
$\lambda_2 = 0.005 - 0.109$	$\lambda_2 = 0.019 - 0.075$

Table 6. Estimation of failure rates of two components across systems.

Some of the areas where meta-analysis has been applied in time series and non-time series are:

- * Combining data from various sources in time and space and linking data across systems (possible in agriculture, medicine, molecular biology, and electronics)

- * Surveys repeated over time with equal or unequal intervals for such items as price elasticity, wage rates, labor force, opinion polls, and impact of high yielding varieties in agriculture in different regions
- * Industrial production estimation, seasonal adjustments, econometric forecasting, using trend analysis
- * Estimating and forecasting failure rates in systems
- * Short term forecasting for power systems and water flow in interconnected systems as in irrigation projects
- * Estimation of missing observations in economic time series using back-shift and forward-shift
- * Reducing uncertainties in short term projects
- * Quality control
- * Signal processing

There is progress in the procedures used to test for biases in publication, to validate the underlying model, to develop a stopping rule for cumulative meta-analysis as opposed to sequential analysis, and to investigate the use of few large studies versus many small studies. Semi-parametric methods are available for this.

Advantages of Kalman Filter in Agriculture and Related Fields

Some main advantages of Kalman filter in agriculture and other fields are:

- * Its recursive nature is ideal for real time and on-line applications.
- * It has attractive optimality properties such as minimum variance unbiased estimates, also maximum likelihood, of state parameters.
- * Its has flexibility to accommodate and validate many time series models and allowing them to be expressed as state-space models.
- * It can detect abrupt system changes such as might be experienced in electrocardiograms (EKG), missile tracking, production runs, etc.
- * It has versatility to link across systems for system reliability.
- * Extended Kalman filter can be used with filtering and interval smoothing for trend level and seasonal effects.
- * The influence of any observation on the system can be obtained via validation estimates of the smoothing parameter.
- * The assumptions of normality and uncorrelated errors which are required for a Bayes analysis may be weakened for Kalman filter.

Research progress can be accelerated through use of appropriate and efficient statistical designs and analyses. This means that biometricians and statisticians need to be involved from planning through the analysis phase of an experiment. Model selection and modification will depend upon the material, conditions and goals of the experiment and real world applications (Rao, 1982; Gnanadesikan and Ketting, 1984). Such statistical procedures as non-parametric estimation, meta-analysis, analysis of variance, regression, multivariate analysis, etc. may be involved, and many researchers will require

help in using and interpreting the outputs. Multi-dimensional scaling, principal coordinate analysis, and other multivariate techniques as well as simulation studies may be useful in interpreting molecular biology and biotechnology data and monitoring size and rate of environmental changes. Data reduction, quality, and cross-validation are necessary components of any research project. There may be intensified use of non-parametric method, particularly for spatial data, and non-parametric regression for restoring functions with noisy observations. Kernel discriminant analysis will find use in frontier sciences like biotechnology. Jackknife and bootstrap procedures are useful to reduce biases, estimate misclassifications, and obtain better discrimination with small samples. Outliers need to be detected and possibly removed in order to prevent distortions of a data set.

There are several areas in Statistics that require extension and development if the needs of researchers are to be satisfied. Present multivariate techniques for the most part are inappropriate for the study of agricultural systems such as inter cropping (Federer and Murty, 1987). Sensitivity analysis and predictive validation, both retrospective and prospective, are needed for modeling assumptions and homogeneity across studies. Model formulation, choice of control variables, and Kalman filter gain are needed for biology, environmental medical and space research. The study of time trends and the use of Kalman filter to construct a weighting function with more emphasis on recent results (batches/experiments) are needed. The application of hierarchical, fixed, mixed and random modeling concepts need more study and application in all fields of inquiry. Better interdisciplinary approaches and efforts to match information from multiple sources and forecasting will be beneficial in advancing knowledge.

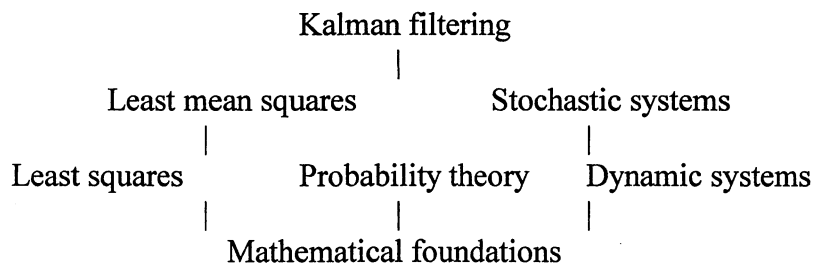
Some conclusions that may be drawn are:

- * Efficient statistical methods are available for analyzing unbalanced data sets without discarding information on controls or checks
- * Combined analysis of trials as in plant breeding using a non-parametric approach is possible and needs to be compared with other approaches
- * Rational clustering of sites in multi-location experiments and testing stability of clusters is possible in an optimal manner
- * The study of interaction between site clusters and genotypes is possible and informative
- * The classification of sites and genotypes by two-stage multi-sample cluster analysis and regularized discriminant analysis may be more useful than empirical grouping based solely on soil data
- * A combination of several methods may be necessary
- * Meaningful use and interpretation is possible only with constant interaction between the statistician and experimenter
- * Unknowledgeable use of software packages (basically black boxes) will result in misapplication and misuse of data and results
- * Theoretical and applied statistical research needs to be more directed toward the real world requirements of users of statistical procedures
- * Kalman filter is a useful procedure in a time series as well as in a non-time time series framework and needs to be used in more agriculture as is being done with clinical studies

NEW DEVELOPMENTS IN KALMAN FILTER: THEORY AND APPLICATION

During the past decade, in addition to the modification of Kalman filter like generalizations of extended Kalman filter, several new developments have taken place in theory and practice. The areas of development include modeling, estimation of bias and its correction, heteroscedasticity in combining data, robustness and sensitivity analysis, new types of filters like integration based Kalman filters, and smoothing algorithms. To appreciate these developments, the fundamental concepts of Kalman filter and the defining of the terms like Kalman filter gain in terms of dynamic linear models is given in the introductory part of this section on new developments. Several reports on new applications in nuclear instruments and methods, high energy physics, and signalling have been published but not referred to here being not immediately relevant to biological experiments. The above reports are summarized in the Jones and Tompkins (1998) publication "A Physicist's Guide to Kalman Filters", World Wide Web, March, 1998, PDF(1.0M/PS(2)). Some of the general problems in meta-analysis needing future work are given by Olkin (1992), Mosteller and Chalmers (1992), Dear and Begg (1992), Hedges (1992), and Gurevitch and Hedges (1993).

Fundamental concepts in Kalman filter--Kalman filter is a special case of time series models called dynamic linear models, DLM, and a part of another discipline, modern control theory. The essential subjects forming the foundations for Kalman filter theory are given figuratively by Grewal and Andrews (1993) as outlined by Kalman (1960) and Kalman and Bucy (1961). These fundamental concepts of Kalman filter are:



Kalman filtering may be employed to carry out a version of model mixing strategy like that moving over time from one dynamic state to another among a finite set of such states as monitoring in transplants in medical practice. Forecasts can be made prospectively by model-mixing approaches, and it remedies the predictive uncertainty in decision making. Kalman filter solved the data fusion problem such as combining radar data with inertial sensor data to arrive at an overall estimate of aircraft trajectory and is equally useful for several similar biological problems. Although Kalman filtering was derived for a linear problem, it may be applied successfully to many non-linear problems. Discrete time and continuous time estimators can be applied to non-linear problems. These extensions use partial derivatives as linear approximations of non-linear relations. The approach of evaluating these partial derivatives at the estimated value of state variables is called Extended Kalman Filter. Thus, it requires linearization of plant and observation

equations. Non-linear filtering using higher order approximations of the filter equations has also been done for quadratic non-linearities and terms through third order. A general approach of non-linear stochastic differential equations was also developed resulting in a stochastic partial differential equation describing the evolution over time of the probability distribution over the state-space of the dynamic system being studied (Grewal and Andrews, 1993).

Kalman gain (K_t) is the adaptive coefficient in a dynamic linear model and is used in variance update, i.e., $C_t = K_t V_t$ and posterior mean m_t for μ_t is a simple weighted average of the most recent observation Y_t and that observed prior mean m_{t-1} where $m_t = K_t Y_t + (1 - K_t) m_{t-1}$. It is defined as $K_t = (C_{t-1} + W_t) / (C_{t-1} + W_t + V_t)$ where V_t is observational variability, W_t is dynamic variability, and taking μ_0 distributed as $N(m, C_0)$ with m_0 and C_0 assumed known.

Smoothing-- A smoother estimates the state of a system at time t utilizing the measurements made before and after time t . Thus, they use observations beyond the time that the state of the system is to be estimated. The accuracy of the smoother is higher than that of a filter because it uses more measurements for the estimate. The three types of smoothers are fixed interval smoothers using all the measurements after a fixed interval, fixed-point smoothers which estimate the system state at a fixed time in the past given the measurements up to the current time, and fixed-lag smoothers which estimate the system state at the fixed-time interval lagging the time of the current measurement (Grewal and Andrews, 1993).

Bias and Spurious Trends and Correction--Presence of publication bias in meta-analysis and adjustment for publication bias using a Bayesian approach was shown in passive smoking data (Dear and Begg, 1992; Givens, Smith, and Tweedie, 1997). Estimation was based on a data-augmentation principle within a hierarchical model, and the number of outcomes of unobserved studies was simulated using Gibbs sampling which showed relative risk was lowered from 1.16 to 1.10 after adjustment for publication bias and could be of relevance in Kalman filter analysis also. Devising a method of handling small collections is suggested as bias will be more in small samples. Response surface exploration in a least squares framework is recommended for unbiased estimation of effects and conditional on ideal study design characteristics (Vanhonaker, 1996) and two significant results which are suspect were identified. Some general combining rules and diagnostic measures to obtain best linear unbiased estimates (BLUE's) of missing values in time series forecasting and recursive estimation using Kalman filter were illustrated by Pena (1997). Treating estimation of a parameter or the forecast of a random variable as a process of combining information, he showed some insights in the robustness properties of some statistical procedures like Kalman filter. In the analysis of trends, autocorrelation of survey errors in routinely used designs induce spurious trends that need to be separated from the trend of population values. This could be done using Kalman filtering that combines the model holding for the population values with the model implied for the survey errors (Pfeffermann, Feder, and Signorelli, 1998) and they demonstrated this using the Australian Labour Force Survey.

Modelling Modifications--Some developments in Kalman filtering are both theoretical and applied in specific situations. Dynamic versions of the model mixing approach are suitable for many applications and forecasts are made prospectively by using

this approach as it remedies predictive uncertainty (Gaver *et al.*, 1992). Models for incorporating historic controls into meta-analysis are also outlined by Beggs and Pilote, (1991). Model-based Gaussian and non-Gaussian clustering was also developed (Banfield and Raftery, 1993). Advances in the techniques and technology of Kalman filters including model-mixing was outlined by Leondes (1982). Adaptive modeling based on the design and the optimization of recursive algorithms by tracking the parameters of time-varying dynamic models via Kalman filter led to significant improvement in forecasting (Leondes, 1982; Grillenzoni, 1994). Adaptive control in the presence of time-varying parameters as in Kalman filter could be specified in an econometric model to discriminate among various policies through the simple use of dummy variables (Tucci, 1997). He showed that Kalman filter permitted a neat application of Bayesian learning within the framework of quadratic objective functions to update optimal forecasting rules. Use of the mixed effects model and models for incorporating historic controls (Beggs and Pilote, 1991) are suggested for amalgamating data such as dose-response relationship (Verducci, 1988) with a Bayesian approach. Empirical Bayes approach is to be developed for biologically motivated stochastic models in situations with added explanatory variables and development of Bayesian methods to quantify the risk (Gaver *et al.*, 1992). Non-parametric Bayesian approach is also recommended in some cases (Tiwari, 1994). Development of random effects models for non-Gaussian data and new statistical methodology on matching models as in DNA profiling and basic studies in matching information for multiple series is recommended (Gaver *et al.*, 1992).

Non-linear filtering of a discrete hybrid stochastic system in an interactive multiple model with small observation noise and the system dynamics is coupled with a finite state unknown parameter was outlined which showed that if the switching is completely observable, the extended Kalman filter becomes a special case of this filter (Kannan and Zhang, 1998). A new approach to state estimation and prediction using a fuzzy supported extended Kalman filter to select a subset of models led to a pronounced improvement in state estimation in more complicated bio-processes as in beer fermentation (Simutis, Havelik, and Lubbert, 1992). With all these selected models, estimation of the current process state was done using extended Kalman filter. Then, all individual results were combined to obtain an effective state estimate using fuzzy reasoning techniques. A posterior mode estimate as by extended Kalman filter and smoother for multivariate dynamic generalized model was illustrated by Fahrmeir (1992).

A generalized state-space model and its application to a set of extended Kalman filter equations was outlined with an example of a more complicated non-linear trend model for financial time series (Papanastassiou, 1992). Bayesian analysis of bi-linear time series model as compared to linear time series was illustrated with two real data sets on gasoline price changes and forecasting (Chen, 1992a) and compared with a linear time series with sun-spot data (Chen, 1992b).

A spectral approach to non-linear filtering based on Cameron-Martin version of Wiener chaos expansion was developed which gave rise to new numerical schemes for non-linear filtering. The algorithm separates the computation involving observations from those dealing only with the system parameters (Lototsky, Mikulevicius, and Rozowskic, 1997).

Heteroscedasticity and Heterogeneity and Kalman Filter -The importance of detecting and correcting for heteroscedasticity in combining data in general including state-space models is emphasized by Piepho (1995) for yield trial data. REML method to estimate variance components where heterogeneity between trials is considered by Frensham, Cullis, and Verbyla (1997) where Kalman filter can be used to combine data. Bayesian approach minimized variability across experiments (Verducci, 1988) in dose-response curves, even in the presence of heterogeneity. Bayesian approach as Kalman filter was suggested to investigate the underlying risk as a source of heterogeneity in meta-analysis of clinical studies (Thompson, 1997). Formulation of a prior distribution with a Bayesian approach for the heterogeneity is suggested in meta-analysis by Higgins (1996).

Grouped random effects models for meta-analysis were presented by Larose (1997) by specifying diffuse proper prior and hyper-prior distributions to observe posterior probability. The sensitivity of the posterior to choice of prior was analyzed by him. A random effects combined analysis within a sequence framework is recommended by designing a protocol and procedures like Kalman filter that incorporate prospectively the elements that are usually necessary in a large well-designed trial (Flather, 1997). Monte Carlo simulation studies in cumulative meta-analysis is suggested by Berkey (1996). Possible use of Kalman filter is to be explored and compared with the Monte Carlo method. Longitudinal data analysis using Kalman filter when each subject is observed at different unequally spaced points with large number of subjects and small number of observations in time showed utility of Kalman filter (Jones and Ackerson, 1990; Jones and Boadi-Boateng, 1991). A state-space was applied to robust regression with discontinuities and to a time series change point (Carter and Kohn, 1996).

New Uses of Filters--New finite dimensional filters are derived which estimate integrals and stochastic integrals of the moments of the state variable and were developed for linear Gaussian dynamics. In the existing literature, parameter estimation of linear Gaussian models via the EM algorithm and filter estimates of the above via a Kalman filter are computed which require large memory. The new procedure provided Maximum Likelihood estimates of the parameters in the dynamics of the Kalman filter (Elliot and Krishna Murthy, 1997). Better M-sequence design for parameter identification using the degree of observability and useful control system design was outlined with improved identification than extended Kalman filter which may give biased estimates due to inaccurate specification of the system functions and noise characteristics (Ohtuska, 1998).

Algorithms for a new class of filters called risk-sensitive (RS) have been developed. The RS filter differs from a conditional mean estimator like Kalman filter and is either risk-prone or risk-averse depending on the sign of a scalar θ which appears in the function. All processes in RS filter retain the Gaussian characteristic. Their departures from Kalman filter rests on the fact that many of the statistical processes explicitly depend on the measurement history and subject to the measurement realization (Banavar and Speyer, 1998).

Taplin (1998) described a new class of Kalman filters called Reverse Kalman filter which complements the conventional Kalman filter recursion for the distribution of the state given the past data. Reverse Kalman filter calculates the distribution of the state given the future data in the state space model. The utility of the two recursions is that

they can be combined efficiently to calculate leave-k-out diagnostics or the likelihood under a model incorporating a patch of anomalies in the series.

In sequential on-line filtering, useful in ecological studies, highly skewed observations and frequent presence of outliers makes the classical Kalman filter not always appropriate. Therefore for on-line filtering, integration-based Kalman filter is applied which leads to estimates of state vector and also dynamic estimates of unknown hyper-parameters (Schmalter, 1994).

A generalized predictive control for joint state and parameter estimation was applied to control estimated biomass in a bioreactor to follow pre-defined trajectories while the biomass was estimated using an extended Kalman filter from on-line measurements and culture volume as extended Kalman filter was not fully satisfactory (Shimitzu, 1993).

A new filtering method call Polynomial filtering for linear discrete time non-Gaussian systems that generalizes a previous result concerning quadratic filtering is developed. A recursive v^{th} order polynomial estimate of finite memory Δ is obtained by defining a suitable extended state which allows the solution of the filtering problem via Kalman linear scheme (Carvetta, Germani, and Raimondi, 1996). The optimal linear estimate of the extended state with respect to the extended observations agreed with the optimal polynomial estimate. Agro-Menyang (1996) used Kalman filter for estimation price levels and the conditional estimating of the stochastic discount factor for pricing kernel and separating the risk and time adjustments. He showed Kalman filter's utility in option pricing, forecasting, and corporate financing by characterizing the time series behavior of the pricing kernel estimates and suggested further refinement.

A two filter adaptive algorithm to combine accurate estimation and rapid adaptation of abrupt frequency changes from noisy sinusoid signals was proposed by Ta-Hsin Li and Kedem (1998). Instead of a single filter, two adaptive filters, a detection filter and a tracking filter, are applied which provide for detection of abrupt changes in the time-varying frequency. This approach is suitable for tracking time varying frequencies that can be modeled as piecewise constant functions of time.

A generalization of extended Kalman filter and smoother for conditionally Gaussian observations in multivariate dynamic generalized models was done for approximate posterior mode estimation which provided good estimates for three data sets (Fahrmeir, 1992). Extended Kalman can be refined further using Mahalanobis Distance Hough transform in multivariate models. Composite vector estimator derived by weighting inversely proportional to variance has been derived (Gregoire and Walters, 1988); the resulting estimator is identical to the single-update Kalman filter. Optimization of an adaptive filter based on information theory can also be done.

Spatial non-homogeneous Kalman filter and its use in image processing is given by Sun, Wang, Xu, and Xu (1996). Research on asymptotic properties of data assimilation algorithm based on Kalman filter is required. Similarly, work is needed on the use of Kalman filter for inference in models with unknown noise distributions.

Robustness and Sensitivity Analysis--Although Kalman filter is useful for filtering and prediction with dynamic linear and non-linear systems, the procedure becomes expensive if the number of state variables is large as in aero-dynamics, medicine, and seismology. Hence, Kalman filter robustification based on M-estimation principle by

using square root formulation of Kalman filter was developed to reduce the costs of $O(n)$ operation for each state update (Romera and Cipra, 1995). This approach tackles the presence of outliers and non-Gaussian distributions in the dynamic linear systems. Procedures for the robustification of Kalman filter were given by Meinhold and Singpurwalla (1989). Corrections of Kalman filter in the presence of outliers is easily accomplished.

An alternative approach to modeling using Bayesian forecasting seasonal time series is outlined using Bayesian auto-regression tradition involving two steps. First, several parameter vectors all having equal prior probability are passed through a given data set using Kalman filter. A sensitivity analysis is then performed to assess how the measure of fit changes when certain parameters are fixed or features of the prior are altered. Secondly, using the parameter vector generating the best fit for the prior to the data, forecasting is done out of sample (Canova, 1992).

Adding an additive term of instability can result in an improved extended Kalman filter. Testing for parameter stability is important in Kalman filter applications, particularly economic. The basic tests for heteroscedasticity are Breusen-Pagan test, fluctuation test, Ljung-Box test, and the Goldfield-Quandt tests described by Wells (1996). All tests do not reach the same conclusions. This can be expected given the length of the data in the samples and care should be exercised in conclusions, particularly if the coefficients of the model are variable. In the estimates of the models, better estimates may be obtained by splitting the sample into two or more sub-samples before estimation.

Smoothing--Kalman smoother equations are based on Kalman filter estimates. A forward pass on the data yields the Kalman filter estimate \hat{x}_n and a backward pass which uses only the latter yields the smoother estimate as explained earlier.

An elegant procedure on fixed interval smoothing recursion for Gaussian models is presented by Kohn and Ansley (1987). Filtering and smoothing formulas apply more generally to non-linear models also. A new formula for the covariance between smoothed estimates at any two points in time is given by De-Jong and Mackinnon (1988). Fahrmeir (1992) provided a generalization of the extended Kalman filter and smoother for conditionally Gaussian observations for approximate posterior mode estimation. Fahrmeir and Wagenpfal (1996) presented a generalized Kalman filter and smoother for estimating hazard functions and time-varying effects in discrete duration and competing risk models and extended the approach to models with multiple terminating events. They developed a numerically efficient Fisher's scoring smoothing algorithm obtained by extending iterative Kalman filter techniques to multi-categorical time series. The smoothing algorithms can be derived as posterior mode estimators. Iterating working Kalman filter and smoother, an interactivity weighted Kalman filter and smoother is obtained. Tanizaki and Mariano (1994) did prediction filtering and smoothing in non-linear and non-normal cases using Monte Carlo integration with less bias than extended Kalman filter and illustrated the same with an example of retail sales estimates and forecasting. Seasonal trend component of business competition was estimated by Kalman filter and fixed interval smoothing with Kalman filter being superior to Monte Carlo filter and smoother (Yada and Kitagawa, 1995). Similarly, Jain (1995) showed that Kalman

filtering and smoothing was superior to ARIMA model-based methods to estimate all the unobserved components with data on unemployment of male workers.

Updating in time series analysis is called exponential smoothing. The more uncertain the state, the greater the weight of the current observation in the revised or updated estimate of the state. Conversely, the more uncertain the observation is, the smaller is the weight applied to the current observation (Wells, 1996). Exponential smoothing as a technique can be used to divide consumption data into a temporary and a permanent part. Permanent part would correspond to the estimated state variable, and the temporary component would be the residual formed by subtracting the estimated value from the actual observation.

Spall (1988) describes the versatility of Kalman filter and its possible use in non-linear and non-Gaussian modeling, estimation including robust methods, system identification, fault detection, and isolation (detecting abrupt changes in models) and intervention analysis. Its utility over traditional time series is in many real world applications. Fixed-interval smoothing is the most useful in time series with trend and seasonal components as in meteorology, econometrics, oceanography, and agricultural and animal production as the state-space model includes local polynomial trend. Smoothness prior modeling methods in time series are extensively applicable. Kalman filter algorithms are the means for the realization of smoothness prior time series modeling.

PART II - KALMAN FILTER APPLICATIONS

INTRODUCTION

Kalman filter is considered by Grewal and Andrews (1993) to be one of the greatest discoveries in statistical estimation history in the twentieth century. Statistically speaking, it is an estimator for the linear-quadratic Gaussian problem of estimating the instantaneous state of a linear dynamic system perturbed by Gaussian white noise. Measurements linearly related to the state of nature but corrupted by Gaussian white noise may be used. Kalman filter provides the means for inferring the missing informations from indirect measurements in cases of variables which are not possible to measure. Its diverse applications to study the control of complex dynamic systems include tracking space crafts, ship navigation, predicting future course of dynamic systems like dynamic flow of rivers, flood predictions, process control in industry, prices of traded commodities, growth of forests, environmental monitoring, abundance of aquatic organisms, biotechnological processes, analytical chemical procedures, epidemiology of pathogens and pests, macro-economic trends in agriculture and animal production, clinical monitoring, and social concerns studies. The principal uses of Kalman filter has been in control theory, predictive design and estimation, and control systems. Applications of Kalman filter theory have resulted in the use of extended Kalman filter for use in non-linear problems, estimation of missing data, and both forward and backward estimation and smoothing in the analysis of trends. As it combines prior information with that contained in the current data to produce estimates and/or

predictions that fully incorporate both sources of information, Kalman filter is a successful procedure for combining information and data. In spite of its demonstrated utility in biological systems, its use in agriculture, and plant breeding in particular, has been minimal and insufficient.

KALMAN FILTER IN FORESTRY

Kalman filter may be used for measuring growth in perennial crops and forests describing crown form and canopy structure, development of disease in stands, sustainability of the system, trend monitoring in land/forest cover, tree growth, and weather relations and estimation of production. The monitoring of time trends using tree ring data for each year in slash pines using Kalman filter was done by Van Deusen (1989, 1991) in which evidence of a 20-year recurring blight was obtained. Evaluation of height-age relationship in pines using Kalman filter and detection of variation in growth patterns within regions resulted in improved prediction for smaller regions and provided a good forest inventory system. Kalman filter was useful in the prediction based on tree ring analysis (Walters, Burkhart, Reynolds and Gregoire, 1991; Visser, 1986; Van Deusen, 1988, 1991; Groton, Eagar, Cook, Ord, Dear, and Taylor, 1988). Discrete Kalman filter was adapted to estimate regression models with time varying coefficients which provided a functional relationship between tree growth and weather (34 meteorological variables) and is suitable for dealing with both gradual and sudden changes in response (Visser and Molenaar, 1990; Van Duesen 1991; Pan and Raynal, 1995; Yude and Dudley, 1995) and predicting tree response to future climatic change. Growth signals indicative of climatic effects were used to construct growth-climate models using Kalman filter which revealed different conifer species have different responses to future climate changes. Evaluation of long-term climatic change of pollution impact on tree growth was done by using and demonstrating the value of tree ring data (Van Deusen, 1987, 1990, 1991). Prediction of future growth and weighting of missing data was achieved using discrete Kalman filter by Visser and Molenaar (1990, 1992) and Le Blanc (1993). Hourly measurements of xylem sap velocity in oak and fir was used to assess influence of atmosphere and air pollution by SO_2 , NO , NO_2 , and O_2 and use of Kalman filter model to detect the influence of gaseous pollutants on weather (Visser, Neppert, Van-Wakeren, and Vaesson, 1988).

Using Kalman filter conditional means and conditional covariance of time inventories were generated which provided a management decision making process by Dixon and Howitt (1979) which helped estimate the unknown time inventories. Tree ring series analysis by Kalman filter from long term growth experiments are suited to intervention analysis in Norway spruce (Van den Brackel, 1996; Makinen, 1997).

Monitoring growth of American beech affected by beech bark disease in Maine was done using Kalman filter by Gave and Houston (1996). The Kalman filter approach was employed to analyze the yearly time dependent mean differences between paired susceptible and resistant tree ring widths using simple structural time series models in state-space form. In one site, moderate maritime climate accounted for the 34 year difference in the onset of decline of trees and harsher winter conditions responsible for the 6 year difference in decline onset in the second site.

Kalman filter was used in the construction of site-index equation to describe height-growth patterns for black spruce. The advantage of Kalman filter over over Richard's growth model being that Kalman filter may be used as base-age invariant site index model (Wang-Yongne and Bijan, 1994).

Kalman filter was used to evaluate temporal variation in associations between physiologically based climate indices and radial growth of black oak which revealed under estimated pollutions in earlier studies (Leblanc-David, 1993; Cook and Johnson, 1989). Mathematical modeling via Kalman filter of forest ecosystems such as tree stability, competition, forest dynamics, and nitrogen cycling was done successfully in Germany (Franke and Roeder, 1991).

Demand for lumber is estimated by USDA Forest Service using a forecasting model assuming that demand coefficients are constant over time. The use of a Kalman filter model indicated that the demand was generally stable in all regions except the Southeast (Seldon and Boyd, 1991). Estimates of elasticity for softwood lumber during 1950-1987 in the USA has been obtained using the Kalman filter algorithm (Adams, Boyd, and Angle, 1992). A significant declining trend in elasticity was found for both in change in consumption and in combined movement of the elasticity series.

A significant air pollution effect on tree ring development was not found to be present in Norway spruce for the different canopy and vitality classes using the discrete Kalman filter method. However, these classes were influenced by precipitation during the growing season (Van den Brakel and Visser, 1996).

KALMAN FILTER IN HYDROLOGY

Kalman filter has been used extensively in hydrology. Studies have included such things as estimation of missing hydrological data, estimates of aquifer parameters from pumping set data, estimates of ground water flow problems due to soil properties, estimates of ground water recharge and its contamination, forecasts of ground water levels in aquifer, estimation of catchment rainfall and run-off, flood forecasts, and water quality in rivers.

Use of Kalman filtering has improved accuracy in the estimation of the first missing data including peak flow. The multivariate regression models are calibrated recursively on available data preceding and following the period of missing data (Bennis, Barrada, and Kang, 1997). Estimation of organic content concentrations in Chesapeake tributaries was performed by Kalman filtering, which more often reflected the true concentration from measured estimates (Godfrey and Foster, 1996). Predictions were obtained for the months where the actual concentrations were unknown. For known concentrations, a filtered smoothed value was also obtained that reflected the correction due to the estimated noise in the measurement and environmental system.

Kalman filter algorithm is used to forecast ground water levels in the upper Floridian aquifer (Graham and Tankersley, 1993). The algorithm processed historical and currently available head measurements to make optimal predictions of future head levels over 554 wells from measurements obtained monthly from a subset of 20 wells and semi-annually from the remaining wells. Estimation of parameters in ground water flow problems was done using a Kalman filter algorithm in the Netherlands (Geer and Kloet,

1985). Automatic control of two different eight-pool irrigation canals and the controller were tested using Kalman filter on a full non-linear model and the estimates were stable, robust, and precise (Malaterre and Pilote, 1998). Ground water measurement in the Netherlands, described by one lumped parameter value of soil properties, was satisfactorily estimated by the Kalman filter algorithm (Geer and Kloet, 1989). Adaptive pumping test-data analyzed by Kalman filtering gave good aquifer parameter estimates (Sen, 1984).

Extended Kalman filter was applied to steady state ground water flow test problems for estimation of transmissivity values and zonation which improved accuracy and stability of results in the USA (Epstein and Dougherty, 1996). River flow forecasting model via Kalman filter was developed using rainfall and run-off information which gave satisfactory forecasts (Dimopoulos, Lek, and Lauga, 1996). Run-off characteristics of small mountainous basins analyzed by Kalman filter gave a good fit in Japan (Yasuda, Takuma, and Sekikawa, 1994). Kalman filtering gave accurate prediction of rainfall and flood forecasting in Japan (Kadoya and Tanakamaru, 1995; Islam, Nagai, and Yomoto, 1994) and in other countries such as New Zealand, Korea, and Denmark (Bidwell and Griffith, 1994; Awwad *et al.*, 1994; Storm *et al.*, 1988). Containment of streams was measured with Kalman filtering for biochemical oxygen demand data in Medowek Creek in North Carolina (Tiwari, Yang, and Zalkikar, 1996), and nitrate concentrations in a stream in Wales with good fit (Sloan, Jenkins, and Eatherall, 1994). Improved concentration predictions of sub-surface transport were obtained via Kalman filtering in the USA (Graham, McLaughlin, and Ralph, 1989). Similar improvement in estimated concentration of contaminants in ground water was obtained with Kalman filtering (Yu, Heidari, and Wang, 1989) and also in the UK (Whitehead, Beck, and O'Connell, 1981). The relationship between run-off predicting models efficiency and basin aridity was studied in Greece with Kalman filtering which performed better with increasing aridity as compared to water balance model efficiency (Mimikow, Hatjisava, Kouvopoulos, and Anagnostow, 1992).

Estimation of organic content concentrations in Chesapeake tributaries in the USA was done using Kalman filter which more correctly reflected the true concentration from measured estimates (Godfrey and Foster, 1996). Prediction was obtained for the months where the actual concentration was unknown. For known concentrations, a filtered smooth value was also obtained that reflected the correction due to the estimated noise in the measurements and the environment system.

The use of extended Kalman filter for state estimation in biological waste water treatment process in Canada allowed on-line tracking of process variables which are not directly measurable (Jones, MacGregor, and Murphy, 1990). Ratio of rainfall to run-off and the length of overland flow was calculated using Kalman filter and the measured values were similar to calculated values (Ichihara, 1987). Real time flood forecasting by coupling meteorological and hydorological models was done using extended Kalman filter and was useful in real time flood and flash flood forecasting (Georgakakos, 1986). Estimation of effective rainfall and real time discharge forecasting by adaptive Kalman filter provided approximately corrected estimates (Kitanidis and Bras, 1980). Extended Kalman filter was used in the urban run-off pollution modeling with rainfall as the

primary input and appeared to be of promise for pollution assessment control (Grum, 1998).

KALMAN FILTER IN FISHERIES

Kalman filter has been utilized in aquatic systems such as fisheries for estimation in production models, abundance, biological dynamics, population parameters related to commercial harvest, prediction of stockage, productivity of sub-populations, role of environment on the population dynamics, estimation of size, productivity, and stock biomass of fish stocks, and estimation of growth and fishing mortality.

A state-space representation of a length structured population under commercial harvest is outlined by Sullivan (1992). Kalman filter is used to estimate the underlying system parameters of fish production such as fishing mortality, selectivity, and initial population abundance, with commercial harvest representing the observations taken on the catch and categorization by length class on the population (Sullivan, 1992). A method of analyzing catch data and its relation with stock abundance using Kalman filter was done for predicting catch per unit effort (Reed and Simons, 1996). Use of data on relative abundance and time series of catch for six sub-populations of fish in the Northwest Atlantic was done utilizing Kalman filter to predict stock size and response to external influences including different management policies in Canada (Collie and Walters, 1991). The estimates were consistent with the observed data. Estimation of population sizes from catch-at-age data using a combination of Kalman filter and the EM algorithm was done (Mendelsohn, 1988), which allowed for multiple sources of observations on different time scales. The study resulted in a new parameterization for both recruitment and fishing mortality based on smoothness priors.

Kalman filter was used for production modeling to estimate the size and productivity of fish stocks from a time series of catches, and relative abundance indices in Alaska. Kalman filter approach was superior to other models and some simulation studies (Kimura, Balsiger, and Ito, 1996). The possibility of improved estimation of growth and fishing mortality in Iceland from catch-at-length data using Kalman filter was shown (Gudmundsson, 1995). Stock assessment and biomass estimation from catch and effort data of cod fish in the North Sea using Kalman filter approach revealed its advantage over traditional models in the UK (Freeman and Kirkwood, 1995).

Regulation of the environment and management adjustment was practiced using Kalman filter for making short term prediction of the daily dissolved oxygen consumption rate data of a shrimp pond in Taiwan (Shang and Liago, 1996). The results demonstrated that Kalman filter was efficient in detecting abrupt abnormalities of the biosystem and detected three transient and two steady state responses in summer. As most fisheries literature include models with variable dynamics and no observation error, deterministic dynamics and observations subject to measurement error, and combined dynamic and measurement availability, a comparison is made of the two paradigms, i. e., Kalman filter and errors in variable, EV, in parameter estimation of the biological dynamics of an underlying stock (Schnute, 1994) with a simple example of catch data and a complex example of catch-at-age data. The utility of the Kalman filter approach was demonstrated in estimating rational priorities for future data collection.

KALMAN FILTER IN ENVIRONMENTAL SCIENCE

Kalman filtering has application in environmental sciences in the areas of measuring atmospheric pollution, pollution in aquatic systems, tracking hurricanes, and environmental policy management (Piegorisch *et al.*; Bril, 1995).

A component model for the time series of SO₂ concentration in the atmosphere using a recursive Kalman filter algorithm was constructed on the basis of spatial analysis in Germany (Shlink, Herbarth, and Tetzlaff, 1997). For forecasting, part of the data was used to establish the parameters and another part to test the extrapolation and all data for trend analysis. Kalman filter method was compared with linear extrapolation and found to produce closer correspondence between the predicted and observed values. Kalman filter was used as a feature selection method and classifier of multivariate data of three near infra-red data sets and a pollution data set in Belgium (Wu, Rutan, Baldovin, and Massart, 1996). Kalman filter successfully selected wave lengths which led to very good results with a correct classification rate. Kalman filter was similarly successful for rapid analysis of passive remotely sensed infra-red data on emission of pollutant species SF₆ from the atmosphere (Brown, 1991). Kalman filter was utilized for recursive prediction of emissions and concentrations at various positions, which obey an atmospheric dispersion model (Mulholland, 1989). The discrete Kalman filter provided optimal estimates of all source rates constituting the state vector and will be of valuable aid in interpreting such data sets. A Kalman filter is constructed to estimate the atmospheric CO₂ concentration and an adaptive scheme is used to estimate the steady state Kalman gain matrix. Measured data are then filtered using the Kalman algorithm. The Kalman filter results are shown to reduce the variability of the air-borne fraction of fossil fuel produced atmospheric CO₂ (Surendran and Mulholland, 1986).

The relationship between climate and modern pollen spectras was calculated from 182 sites resulting in an "analog climat". In the second step, an original combination of canonical correlation and principal components extracted the common information from several fossil pollen sequences. In the third step, the signals from the first two steps are merged by Kalman filter into a final reconstruction of climate where the noise is reduced by 37%. A novel method of use of the extended Kalman filter in pollen detection in aquatic systems by the identification of photosynthetic light models for a Danish river demonstrating the general usefulness of the recursive analysis is applied to a time series of environmental data. Kalman filter has also been used for better on-line estimation of activated sludge bulking (Chen and Beck, 1993).

The role of information structure and parameters of an optimal regulatory mechanism is examined by Papakyriazis and Papakyriazis (1998). It was shown that optimal control policy involves the derivations of regulatory controls, measurement controls, and the sequential estimation of inaccessible emissions and concentrations via a Kalman filter approach.

KALMAN FILTER IN BIOTECHNOLOGY

The application of Kalman filtering in biotechnology for the past two decades is in the area of on-line estimation in process control, estimation and prediction of biomass, growth rates, and product formation and enzymatic actions detection of important state variables in bio-processes and control of growth of individual and mixtures of cells (13 publications). The utility of Kalman filter to control bio-processes at their optimum state to control costs is also demonstrated.

On-line control and estimation of biomass concentration in *E. coli* and monitoring growth of *E. coli* on glycerol using Kalman filter on on-line measured data was done by Dubach and Markl (1992). It was found that in biomass concentration up to 50 grams dry weight per liter, the estimation of the process was accurate but at higher concentrations, product formation with brown color was detected and the estimated data diverged from the experimental data as the effect of this product on biomass production was not included in the model. Similarly, on-line estimation of biomass concentration and of the three variable parameters by extended Kalman filter was demonstrated in yeast grown in aerobic conditions on an ethanol substrate (Nahlik and Burianec, 1988). The utility of Kalman filter in estimating biomass concentration, which is difficult to measure was demonstrated. Process control of reactor models of continuous production of ethanol from sucrose with immobilized yeast with the use of Kalman filter to maintain constant concentration in the fermentation effluent was shown (Mandenius, Mattiasson, Axelsson, and Hagander, 1987). Profile control of specific growth rate of yeast using an extended Kalman filter was achieved in a baker's yeast fed-batch culture, which cannot be measured directly (Shimitzu, Shioya, Suga, and Takamatsu, 1989). Optimal production of glutathione was obtained by combining specific growth rate of yeast in fed-batch culture with extended Kalman filter and controlling the optimum profile by manipulating the substrate feed rate in the culture (Shimitzu, Araki, Shioya, and Suga, 1991). As a result, the maximum production of glutathione was successfully accomplished. Similarly, improved state estimations and prediction of alcohol fermentation in beer brewing using extended Kalman filter in complex bioprocesses were achieved (Simitis, Havelik, and Lubbert, 1992).

Utilizing extended Kalman filter, on-line estimation of the specific rate which is directly unmeasurable, was achieved in batch or fed-batch fermentation process (Shimitzu, Takamatsu, and Suga, 1989). Criteria for judging the validity of the estimated value from the observed data are proposed. Based on the proposed criteria, the system equation of the specified growth rate is selected and the initial value of the state variable and a covariance matrix of the system noises was adjusted. Extended Kalman filter using covariance matrix, a constant element when cell concentration is measured directly, has estimated accurately predicted values of specific growth rates than the adaptive extended Kalman filter. Bioreactor control was achieved using an extended Kalman filter to remove the effect of noises from dissolved O_2 concentration measurements (Lee, Hwang, Chang, and Chang 1991) and thus to improve control performance relating two control variables air flow rate and agitation speed with dissolved O_2 concentration. The new Kalman filter algorithm performed better than other algorithms tested.

Extended Kalman filter is used to demonstrate the practical use of Kalman filter under two cultivation modes using glucose and ammonia as observation variables for the on-line estimation of viable cell and monoclonal antibody concentration during batch and fed-batch cultures of OKT-3 hybridoma cell lines (Ghoul, Dardenne, Fontiex, and Marc, 1991). The concentration of viable cells and antibodies was twice as high in the fed-batch as in the batch culture.

Control of mixed culture bio-reactor with competition and external inhibition in the growth of two species in a continuous stirred tank was achieved using extended Kalman filter to estimate the unmeasured states when only a single measurement of the total cell mass is available (Hoo and Kantor, 1986). It was shown that one species is sensitive to an external inhibitor for which both species compete for the same rate-limiting substrate. It was also shown that choosing the dilution rate and the inhibition addition rate as manipulated variables admitted a global linearized transformation which may be used to construct a multivariate feedback controller.

Extended Kalman filter was applied to study enzymatic deactivation in the enzymatic hydrolysis of pre-treated cellulose, replacing an earlier model. This approach allowed the study of enzyme deactivation without additional experiments and within operational conditions (Caminal, Lafuente, Lopez, Poch, and Sala, 1987).

Extended Kalman filter was also applied to a continuous cell culture model for bioassay and fermentation estimates by Svrcek, Elliot, and Zajic (1974) as a useful tool for studying biosynthesis. The competitive growth of two phenotypes, Lew+ and Lew-1, and their product formation with a recombinant yeast strain which contains a shuttle vector with the foreign gene (hepatitis B virus surface antigen) was examined. Extended Kalman filter was used to estimate seven state variables in which three variables such as ratio of plasmid-free cell concentration to plasmid containing cell concentration (σ) expression of human hepatitis B surface antigen (g, CH) and change of working volume in the fermentation (V) are experimentally observable (Shi, Ryu, and Yuan, 1993). Simulation results agreed with the observed, and the method of predicting an optimum profile of the cell growth under different dissolved O_2 concentration was demonstrated. Shimizu (1993) demonstrated the use of Kalman filter and extended Kalman filter to control bioprocesses at their optimal states and thus controlling production costs.

Bacteriological infection in biological samples could be detected to monitor food infection by using a multi-process Kalman filter which provided an on-line estimate of the change point reflecting the growth of the organism (Whittaker and Fruhworth-Schatter, 1993). The procedure was versatile in providing evidence of the change and estimated value of the change point and also provided filtering and smoothing analysis of the dynamic state-space model. The estimated value correlated with the other estimates of the strength of the original infection in the sample. All these instances illustrate the powerful utility of Kalman filter in a frontier science like biotechnology.

KALMAN FILTER IN BIOCHEMISTRY AND ANALYTICAL CHEMISTRY

Kalman filter is used in biochemistry and analytical chemistry for the simultaneous detection and estimation of groups, compounds, dynamics of industrial processes, estimation and differentiation kinetics of isozymes and other compounds,

resolution of complex or overlaps, detection and correction in fluorescence responses, and analysis of ultra-violet (UV) and NMR spectra.

Simultaneous determination of three compounds, i.e., caffeine, propyl-phenazone, and salicylamide, and quantification of the concentration in tablet formation using a rapid reproducible UV spectrophotometric multi-component method based on a Kalman filter was done which agreed well with those obtained by HPLC system. The assay results had low standard deviations (Patsi, Malkki, and Tammilehto, 1992). Similar success was obtained in the simultaneous determination of cobalt, copper, zinc, and cadmium in synthetic samples and environmental samples by spectrophotometry and Kalman filter method (Shi, Li, Xu, Pan, and Wang, 1991). Kalman filter was applied to resolve the overlapped absorption curves of the above complexes, thereby making the simultaneous determinations of the metallic ions possible without tedious pre-treatment. Kalman filter was also used for the simultaneous determination of four amino acids, i. e., tyrosine, cryptophan, phenylalanine, and 3-4 dihydroxyphenylalanine by direct UV spectrophotometry combined with a Kalman filter algorithm, resolving the overlapping UV spectra of the four amino acids (Shi, Xu, Pan, Liu, Gao, Qiam, Nie, and Li, 1990). The results favorably compared with those through target transformation factor analysis method. Similar success with Kalman filter was obtained in the UV spectral analysis of mixtures of solutions of six amino acids (Pan, Xia, Si, Zhang, Shi, and Liu, 1990).

The accuracy, simplicity, speed, and recursive nature of linear and non-linear Kalman filter were demonstrated in the analysis of differential kinetic data to estimate differential reaction rates in a mixture of cortisone and hydrocortisone by reaction with Blue Tetrazolium (Xiong, Velasco, Silva, and Perez, 1991). Kalman filter yielded better results compared to logarithmic extrapolation, single point, and proportional equation methods.

Recursive Kalman filter algorithm was used for parameter estimation for analysis of solution kinetics of bovine intestinal and liver alkaline phosphatase (ALP) isozymes in the presence of a denaturant (Lewis and Rutan, 1991). Successful quantification of the two isozymes in synthetic mixtures was achieved. This method will serve as a basis for the development of electrophoretic separation method for ALP quantification with differential kinetic detection. Extended Kalman filter was also used to estimate initial reaction concentrations and rate constants for rate-based chemical assays employing a second order chemical reaction. The method decreased the computational burden of large data sets (Barker and Brown, 1988).

Data on full fluorescence spectral detection in liquid chromatography in the analysis of polyaromatic hydrocarbons was analyzed using Kalman filter based methods for adaptive subtraction of background responses, swift correction, and linear regression analysis of overlapped responses (Cecil and Rutan, 1990, 1991), which were chromatographically unresolved. Kalman filter was similarly used for the correction of spectral response shifts in overlapped fluorescence spectra of mixtures of polycyclic chromatic hydrocarbons (Cecil and Rutan, 1990).

Kalman filter was used to improve greatly in the throughput of the total analytical procedures reducing analysis time without adversely affecting the accuracy and precision of the chromatographic system (Matsuda, Hayashi, Ishibashi, and Takeda, 1988). Kalman filter method resolved overlapped peaks and quantitatively analyzed complex

chromatograms of diazepam compounds. Similarly, unresolved peaks in HPLC (High-performance liquid chromatography) could be deconvoluted and quantified precisely by Kalman filter such as in naphthalene and diphenyl peaks (Hayashi, Shibasaki, Matsuda, and Uchiyama, 1987). Kalman filter thus eliminated the skillful contrivance of the chromatographic conditions and would be applicable to other chromatographic techniques also. Kalman filter approach resulted in the smallest errors in the separation of two amino acids, glycine and glutamine, with similar fluorescence response with severe chromatographic and spectroscopic overlaps (Rutan and Motley, 1987).

A fuzzy-supported extended Kalman filter provided a new approach in the state estimation and prediction in alcohol formation in beer brewing with estimation and prediction of the time at which the beer malt fermentation process must be stopped and also estimated short term and long term predictions of the process (Simutis, Havelik, and Luebbert, 1992).

Double iterated Kalman filter was used for the calculation of the structure of protein from NMR data with results comparable to two other methods DGEOM and XPLOR used for the calculation of the structure (Liu, Zhao, Altman, and Jardetzky, 1992). Extended Kalman filter was used for extracting maximum information for the data evaluation of analytical signals from Flow-Injection Analysis (FIA) peaks. Kalman filter reduced the measuring dead-time of FIA system and corrected noise by estimation of the off-set of FIA peaks and smoothing of FI signals (Wu and Bellgardt, 1998).

Bacteriological infection in biological systems could be detected to monitor food infection by using a multi-process Kalman filter which provided the on-line estimate of the change point reflecting the growth of the organism (Whittaker and Fruhworth-Schatter, 1993) as mentioned earlier under Biotechnology. The procedure was versatile in providing evidence of the change and the estimated value of the change point but also provided filtering and smoothing analysis of the dynamic state space model. The estimated values correlated well with the other estimates of the strength of the original infection in the sample. Thus, Kalman filter provided an excellent supporting procedure in complex analytical chemistry.

KALMAN FILTER IN MEDICINE

Kalman filter procedures have been applied in medicine in the fields of signal processing (ECG, EEG, signals, etc) monitoring the progress of patients, as in renal transplants, and analysis of complex data as hemodynamics and vision problems, trend detection and analysis of spectra to name a few. Twenty three major references are covered in this review.

Signal processing is done in analyzing ECG, EEG, and EMG using Kalman filter algorithms. A multi-stage Kalman filtering algorithm within a prototype ICM (intelligent cardiovascular monitors) was used in the high level analysis of the patients condition to help detect ECG arrhythmias (Sittig and Cheung, 1992) providing a faster and reliable means of accurately detecting ECG arrhythmias in real time. As availability of the R-R intervals in ECG yield useful information on the various types of arrhythmia, Kalman filter was used to detect the onset of arrhythmia and actually identify the arrhythmia (Woolfson, 1991). Thus Kalman filter has a general application in the study of normal

arrhythmic segments of data by an analysis in the calculation of time varying spectra of the data. Automatic evaluation and classification of pattern of epileptic EEG (Penczek, Grochulski, and Kowalczyk, 1986) indicating a global description of a seizure in terms of Markov-chains. Using a Kalman filter algorithm of the multi-channel case, and its application to a segmenting procedure in the analysis of epileptic EEG, estimation of the information flow between a set of four-channel EEG recordings was done in the structural analysis, using the segments extracted by the Kalman filter (Penczek, Grochulski, Grzyb, and Kowalczyk, 1987).

Kalman filter was used for the reduction of super-imposed noise on EEG tracing and to capture the disturbances of the electromyograph data from 20 patients (Bartoli and Cerutti, 1983). Thus Kalman filter has a marked reduction of noise without distorting the useful information contained in the signal. Kalman filter algorithm was used for optimal adaptation of the signal templates in the detection of the motor unit action potential (MUAP) wave forms in an electromyograph (EMG) for clinical diagnosis and therapy by detecting as many MUAP as possible in a single measurement (Studer, Figueiredo, and Moschytz, 1984). The algorithm was successfully applied to synthetic and real EMG data. Kalman filter was used in echocardiography to quantify aortic regurgitant orifice and volume based measurements of the velocity of the regurgitant jet, aortic systolic flow, systolic and diastolic arterial pressures, a wind kessel arterial model, and a parameter estimation technique (Slordahl, Solbakken, Piene, Angelsen, Rossvoll, and Samstad, 1990). Thus it was possible to quantify regurgitant orifice and volume in patients non-invasively from doppler and blood pressure measurements.

Monitoring breast cancer patients following surgery to detect departures of tumor markers from steady states was accomplished using Kalman filter (Schlain, Lavin, and Hayden, 1992). An approach to monitor serial marker data of large numbers of patients even when the series is short and the data are serially correlated and unequally spaced could be used to recommend appropriate testing intervals. A multi-state Kalman filter algorithm was used for accurate and reliable detection and identification of trends, abrupt changes, and artifacts from multiple physiological data (Sittig and Factor, 1990). Such an analysis illustrates the potential value of Kalman filter in physiological monitoring. Similar application of Kalman filter in heart rate monitoring and detection of change in a time series and predicting future observations was done. Multi-state Kalman filter in medical monitoring provides a powerful method in a variety of settings for time series analysis such as detection of kidney transplant rejection (Gordan, 1986). By using a multi-process Kalman filter to detect changes in trends of plasma creatinine and urea concentrations in renal function, dysfunction was identified by the computer much earlier than by the clinician and in 11 of 1259 days when the clinician did not suspect rejection in addition to the identification of 31 of 32 episodes of deterioration in renal function (Smith and West, 1983; Trimble, West, Knapp, Pownall, and Smith, 1983; Gore and Bradley, 1988). Thus, Kalman filter provided for immediate analysis of incoming results and for timing events either prospectively or retrospectively. Kalman filter was useful in monitoring bone marrow transplants (Grillenzoni, 1994). It was used to estimate two tissue parameters, i.e., reflection coefficient and attenuation coefficient, from the reflected ultrasound images and obtaining information on tissue characterization (Shiima, Ikeda, and Saito, 1987). Kalman filtering is therefore an excellent method for clinical cases.

Kalman filtering was superior to ordinary autoregressive spectral estimation for the computation of running frequency spectra from non-stationary oscillations in a long time series and recognizing rapid changes in the frequencies of oscillations (Skagen, 1988). It was also efficient in filtering noisy aortic flow and pressure wave forms of hemodynamic data as Kalman filter does not require a preliminary identification of the signal generation process (Jetto, 1985). The filtering procedure was fast compared to traditional methods requiring computationally onerous procedures,

Deconvolution of noisy signals and analysis of radio-nuclide angiocardigraphy time series curves in humans was done using Kalman filter (Commenges and Brendal, 1982). This method improved the reliability of results in left-to-right cardiac shunt quantification.

A Kalman filter approach to eye tracking movements to detect and separate smooth pursuit and saccadic eye movements proved highly reliable and could be used for other eye movements such as nystagmus (Sauter, Martin, Renzo, and Vomscheid, 1991). Kalman filtering could also be used in neural networks for vision and image processing (Pentland, 1992).

A state space model, Kalman filter, for HIV epidemic was applied to a homosexual population for estimating the number of susceptible people and AIDs cases, and for estimating the probability of HIV transmission from infective to susceptible people (Tan and Xiang, 1998) and measles in New York State (Kalivianakis, Mous, and Grasman, 1994). The results demonstrated that the estimated number of AIDs incidences agreed closely to the observed indicating the usefulness of the model. Meta-analysis was useful in identifying information process factors that characterize children of developmental coordination disorder (DCD) and that visual spacing processing is implicated in DCD. The use of Kalman filter could be equally helpful (Wilson and McKenzie, 1998). Kalman filter was also applied in the analysis and projection of birth outcomes due to illicit drug use and health (Mocan and Topyan, 1993). It would be interesting to see the application of Kalman filter to the aspirin-heart attack data ((Canner, 1987; Gaver *et al.*, 1992; Higgins, 1996) to see if improved estimates can be obtained.

KALMAN FILTER IN AGRICULTURE

In agriculture, Kalman filter has been applied to a limited extent in areas of estimation and prediction of productivity of crops, forest, animals, poultry, and fisheries, analysis of soil moisture, nutrient and productivity status, monitoring and management of production of crops and animals, dynamics of disease pathogen and pest build-up and control, monitoring of process and quality of food products, trend analysis of production for prediction, effect of climatic changes and predicting supply/elasticity and demand and smoothing in biological and econometric models. It is still to be used in important areas such as plant breeding, pathology and pest control, agronomy and quality control, agricultural chemicals and residue analysis, marketing and pricing, agricultural systems analysis, and sustainability of production systems. Fifty publications on the uses of Kalman filtering were found in the literature.

The feed intake, growth, and beef production by growing beef cattle is difficult to predict and can be projected more precisely from past performances for which Kalman

filter is eminently suitable. It provides an advantage in early prediction and trend estimation with lower bias of prediction for feed intake and weight gain (Oltjen and Owens, 1987). Monitoring milk production and daily yield was done and production predicted using Kalman filter (Van Bebber, Reinsch, Junge, and Kalm, 1997). Kalman filter was used for the estimation of the lactation curve of a dairy cow which allowed the inclusion of prior information on the curve, taking into account the correlation between successive observations (Goodall and Sprevak, 1985). The procedure gave accurate estimates of milk yield at early stages of lactation. Kalman filter was used in the prediction of parameters like daily milk yield, oestrus, and electrical conductivity of the milk (for detecting cows with mastitis) which were close to the observed values. Changes in poultry production responses and food intake were detected using Kalman filtering which also showed trends such as steady state, slope changes, and transient responses which could be monitored daily (Roush, Tomyama, Garnabui, Alfonso, and Cravener, 1992) which allowed for management adjustments for the flock. Similarly, prediction of commercial harvest of fish and estimation of population parameters was achieved using Kalman filter (Pella, 1993) and by developing the conditional likelihood equation needed for estimating the underlying population parameters of a length-structured fish population (Sullivan, 1992). Kalman may be used to study the role of environment on population dynamics of fish (Mendelssohn, 1988). Kalman filter was also used to predict forest inventories and decision making on inventories (Dixon and Howitt, 1979).

Use of Kalman filter in the prediction and trend analysis of perennial crop supply and agricultural products such as corn and alfalfa, was done successfully. In estimating perennial crop supply response, problems arise owing to the lack of data on new plantings, removals, and area in individual age groups. Kalman filter was used to generate parameter estimates as well as estimates of new plantings, removals, and existing acreage by age group (Knapp and Konyar, 1991). The results showed that existing acreage has differential impact on new plantings and removals depending on age in California lucerne production. A growth model using Kalman filter was developed to study the combined effects of CO₂ and sucrose on the growth of alfalfa cuttings in greenhouses/culture from data on plant dry weight, leaf number, and percent root initiation (Tani, Murase, Kiyota, and Honami, 1992). The model indicated that dry weight increase would be greatest with 2,000 ppm CO₂ and no sucrose in the medium. Experimental data agreed well with the predicted data. Kalman filter was used to estimate USA maize yield model which incorporated a stochastic trend term and monthly weather indices (Kaylen and Koroma, 1991) and to estimate future maize yields. The estimate for 1989 was close to the final USDA estimate.

A knowledge of soil moisture dynamics is necessary to provide efficient irrigation management. Kalman filter state-space models were applied using soil water balance and potential evapotranspiration for estimation of temporal and spatial soil water storage with good agreement between simulation and field results (Hanks, 1992). Thus, Kalman filter could be extended to other systems where only a few measurements are available. Estimation and forecasting soil water depletion and crop evapotranspiration using Kalman filter in Colorado allowed irrigation decisions to be made (Aboitiz, Labadie, and Heermann, 1986) and prediction of soil properties (Lin, 1994; Hartum and Loftis, 1987).

Kalman filter was also used to estimate ground water recharge from infiltration of rain water using daily water table levels and rainfall measurements of an unconfined coastal aquifer in Australia (Viswanathan and Evans, 1983). An extended Kalman filter was used to determine hydraulic conductivity functions in field soils (Wendroth, Katul, Parlange, Puente, and Nielsen, 1993). The model provided a description of *in situ* hydraulic conductivity as compared to other methods.

Optimal control theory using Kalman filter was applied to the automatic control of two different eight-pool irrigation channels with stable, robust and precise estimates (Malaterre, 1998). Kalman filter was used to develop a feed-back control algorithm for constant volume control of an irrigation canal with estimation of state variables that were not measured (Reddy, 1994). Remote sensing data of rough surfaces (di-electric constant, surface rms height, and correlation length) was analyzed using Kalman filter to invert the surface parameters which enhanced its potential application to remote sensing of rough surfaces (Chen, Kao, and Tzeng, 1995). Kalman filter was used to monitor plant moisture conditions to predict water status of plant for intelligent water management systems (Murase, Nishiura, and Mitani, 1997). Kalman filter was used in the spectral analysis of solar activity and solar constants useful for utilizing solar energy in agriculture (Tadros and Shaltout, 1989).

Kalman filter neural training algorithm was used in monitoring and simultaneous control of temperature and humidity in a confined space in plant physiology studies on leaf transpiration (Murase, Yamauchi, and Honami, 1992). Kalman filter approach maximized the accuracy of the description of the dynamic system. Monitoring plant growth for the development of plant production systems in space using non-linear relation between plant growth and textural features was done using Kalman filter. The estimated leaf size values agreed with those of measured data in lettuce (Murase, Nishiura, and Honami, 1994). The results indicated that the reflection of light over the population of plants varied with the increase in area covered by green leaves.

Monitoring of fruit quality was done using Kalman filter to estimate fruit parameters of the dynamic behavior of fruit with respect to vibration with random forces (Ikeda, 1990). This had provided non-destructive testing for quality evaluation.

Monitoring of animal systems such as milk quality and animal health using Kalman filter was also reported. Bulk tank somatic cell counts (BTSCC) are used to monitor milk quality and to detect mastitis in dairy herds. Kalman filter was used to reduce noise in BTSCC data, retrospective smoothing of the data, detecting presence of outliers, and abrupt changes in the level of the data and forecast future BTSCC values (Thysen, 1993). Pre- and post-stun EEG's after an electric shock to cattle were analyzed using Kalman filter to detect between animal variation and the amount of local smoothing to adopt for the data requirements (Jones and Petitt, 1992). Tracking and investigating constrictions along lymphatic vessels in guinea pigs was done using Kalman filter for real time tracking of visualized biological events (Beresford, Nesbitt, and Van-Helden, 1993).

Use of Kalman filter in the prediction and estimation of agricultural supplies, prices of produce, insect and pathogen increase, and structural changes in marketing are reported. Application of Kalman filter in agricultural economic forecasting in a variety of situations was demonstrated by Lin, Zhang, and Xin (1996). Prediction of prices of fruit and vegetables on the wholesale market seven days and 30 days ahead was done using

Kalman filter taking into account weather data and social factors which revealed seasonal fluctuations and price and supply fluctuations (Akimoto, Kuroda, Kajita, 1986). Kalman filter provided generalization of exponential smoothing and estimating a variety of econometric models and has several useful statistical properties (Schneider, 1987).

Use of Kalman filter estimator for pest management of *Ceratomyxa trifurcata* on soybean was demonstrated which resulted in substantial gains in estimate efficiency and a reduction in the number of fields required to be sampled (Zavaleta and Dixon, 1982). Several examples relating to pests in agriculture, veterinary, and medicine were cited in the prediction and forecasting of insect pest population fluctuations using Kalman filter (Poole, 1978). Kalman filter was used to resolve the mixture peaks in HPLC chromatograms of two mycotoxins alternanol attenuisol and in extracts of *Alternaria* cultivars on rice and maize (Rotunno, 1992).

Kalman filter was used in tracking oscillations in the deceleration phase of the cardiac cycle by accurate estimation of velocity estimation audio-poppler signals (Talahami and Kitney, 1988) and can be applied to animal pathology. Kalman filter was used in the detection of oestrus and mastitis showing its utility in livestock production (De, Kroeze, Achten, Maatje, and Rossing, 1997). Analysis of market dynamics of supply response and deriving of supply elasticities was done using Kalman filter in Slovenia agriculture and bringing out the role of free market versus state regulation in marketing pork, potato, maize, wheat, and milk (Erjavec and Turk, 1997). Kalman filter facilitated the detection of crucial time periods when certain structural variations in price elasticities can be expected (Rao, 1987; Turk, Erjavec, and Gambelli, 1997). Inelastic and declining demand for soft lumber was observed in the USA by using Kalman filter algorithm (Adams, Boyd, and Angle, 1992). Co-integration analysis and parameter estimation procedure of Kalman filter were applied to examine market integration of Australian and USA beef prices using monthly data from 1972 to 1993 (Diakosavvas, 1995). It was found that Australian and USA beef prices are co-integrated although not fully. Kalman filter approach was adopted to handle parameter variation in a study of structural change in the USA demand elasticity for meat in the 1970's. It was found that the income elasticity for beef demand decreased sharply in the late 1970's (Chavas, 1983).

Kalman filter was useful in fusing two data sources, i.e., odometry and image analysis, to establish the accuracy of an autonomous vehicle working in a field of transplanted cauliflower. The accuracy was sufficient to control the vehicle (Marchant, Hague, and Tillet, 1997).

Extended Kalman filter was utilized to identify photosynthesis-light models for a Danish river using 12 sets of data from a natural system (Cosby, Hornberger, and Kelly, 1984). Kalman filter provided better estimators than bulk data estimation method. The effect of climatic change in France during late quaternary period was reconstructed using Kalman filter. At 13,000 year B.p, the climate reached the modern level of precipitation (Guiot, 1987).

Multi-dimensional Kalman filter in genetic programming was more powerful for automatically generating computer programs with the process of natural selection than multi-dimensional least squares regression problem to manipulate multiple data types (Montana, 1995). The above cases indicate the potential of Kalman filter for agricultural

research and it could be useful in plant breeding, plant protection, and sustainable agricultural systems where it is still to make its appearance.

REFERENCES

- Aboitiz, M., J. W. Labadie, and D. F. Heermann (1986). Stochastic soil moisture estimation and forecasting for irrigated fields. *Water Resources Res.* 22:180-190.
- Abraham, B. and K. Vijayan (1992). Time series analysis for repeated surveys. *Communications in Statistics (Simulation)* 21:893-908.
- Adams, D. M., R. Boyd, and J. Angle (1992). Evaluating the stability of softwood lumber demand elasticity by end-use sector: a stochastic parameter approach. *Forest Sci.* 38:825-841.
- Agoro-Menyang, S. A. (1996). Stochastic discount factor estimation using state-space theory and the Kalman filter: Methods and applications. Ph. D. Thesis, Cornell University, 97 pp.
- Akimoto, K., S. Kuroda, and M. Kajita (1986). Prediction of vegetable and fruit prices on the central wholesale market . III: prediction of next day strawberry prices; IV: prediction of strawberry prices a week and a month ahead. *Res. Bulletin Faculty Agric., Gifu Univ.* 51:71-81, 83-90.
- Awwad, H. M., J. B. Valdes, and P. J. Restrepo (1991). Streamflow forecasting for Han River Basin, Korea. *J. Water Resources Planning Management* 120(5):651-673.
- Banawar, R. N. and J. L. Speyer (1998). Properties of risk sensitive filters/estimators. *IEE Proc. Control Theory and Applications* 145:106-112.
- Banfield, J. D. and A. E. Raftery (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics* 49:803-821.
- Barker, T. Q. and S. D. Brown (1988). Estimation of second-order chemical kinetic parameters by using information-based extended Kalman filtering. *J. Chemometrics* 2:137-154.
- Bartoli, F. and S. Cerutti (1983). An optimal linear filter for the reduction of noise superimposed to the electro? Encephalographic Signal. *Med Eng.* 5:274-280
- Beggs, C. B. and L. Pilote (1991). A model incorporating historic controls into a meta-analysis. *Biometrics* 47:989-907.
- Bennis, S., F. Berrada, and N. Kang (1997). Improving single-variable and multivariable techniques for estimating missing hydrological data. *J. Hydrology* 191:87-105.
- Beresford, S. B., K. V. Nesbitt, and D. F. Van-Helden (1993). Edge detection at multiple locations using a "radar" tracking algorithm as exemplified in isolated guinea pig lymphatic vessels. *J. Neuroscience Methods* 49:69-79.
- Berkey, C. S. (1996). Uncertainty of the time of first significance in random effects cumulative meta-analysis. *Controlled Clinical Trials* 17:357-371.
- Bril, G. (1995). Forecasting hurricane tracks using the Kalman filter. *Env. Mtre.* 6:7-14.
- Brown, R. G. (1983). *Introduction to Random Signal Analysis and Kalman Filter*. Wiley, New York.
- Brown, S. D. (1991). Real-time filtering of data from mobile passive remote IR sensors with principal component models of background. *J. Chemometrics* 5:147-162.

- Caminal, G., J. Lafuente, J. Lopez-Santin, M. Poch, and C. Sola (1987). Application of extended Kalman filter to identification of enzymatic deactivation. *Biotechnology Bioengineering* 29:366-369.
- Canner, P. L. (1987). An overview of six clinical trials of aspirin in coronary heart disease. *Statistics in Medicine* 6:255-263.
- Canova, F. (1992). An alternative approach to modeling and filtering seasonal time series. *J. Business and Economic Statistics* 10:97-108.
- Carlin, B. (1992). Comment on "Hierarchical models for combining information and for meta-analysis". *Bayesian Statistics* 4:336-338.
- Carter, C. K. and R. Kohn (1996). Markov chain Monte Carlo in conditionally Gaussian state space models. *Biometrika* 83:589-601.
- Carvetta, F., A. Germani, and M. Raimondi (1996). Polynomial filtering for linear discrete time non-Gaussian systems. *SIAM J. Control and Optimization* 34:1666-1690.
- Cecil, T. L. and S. C. Rutan (1990). Correction for fluorescence response shifts in polyaromatic hydrocarbon mixtures with an innovations-based Kalman filter method. *Analytical Chemistry* 62:1998-2004.
- Cecil, T. L. and S. C. Rutan (1991). Fluorescence detection in liquid chromatography with an intensified diode-array detector. *J. Chromatography* 556:495-503.
- Chavas, J. P. (1983). Structural changes in the demand for meat. *Amer. J. Agric. Economics* 65:148-153.
- Chen, C. W. S. (1992). Bayesian analysis of bilinear time series models: A Gibbs sampling approach. *Communications in Statistics (Theory and Methods)* 21:3407-3427.
- Chen, J. and M. B. Beck (1993). Modelling, control and on-line estimation of activated sludge bulking. *Water Sci. Technology* 28:249-256.
- Chen, K. S., W. L. Kao, and Y. C. Tzeng (1995). Retrieval of surface parameters using dynamic learning neural network. *Int'l J. Remote Sensing* 16:801-809.
- Collie, J. S. and C. J. Walters (1991). Adaptive management of spatially replicated groundfish populations. *Canadian J. Fish. Aquat. Sci.* 48:1273-1284.
- Commenges, D. and A. J. Brendel (1982). A de-convolution program for processing radio tracer dilution curves. *Comput. Programs Biomed.* 14:271-276.
- Cook, E. R. and A. H. Johnson (1989). Climate change and forest decline: a review of the red spruce case. *Water, Air, and Soil Pollution* 48:127-140.
- Cosby, B. J., G. M. Hornberger, and M. G. Kelly (1984). Identification of photosynthesis light models for aquatic systems 2. Application to a macrophyte dominated stream. *Theory Simulation Model.* 23:1-24.
- Cullis, B., B. Gogel, A. Verbyla, and R. Thompson (1998). Spatial analysis of multi-environment early generation variety trials. *Biometrics* 54:1-18.
- Dear, K. B. G. and C. B. Begg (1992). An approach for assessing publication bias prior to performing a meta-analysis. *Statistical Science* 7:237-245.
- De-Jong, P. and M. J. Mackinnon (1988). Covariance for smoothed estimates in state space models. *Biometrika* 75:601-602.
- De, M., R. M., G. H. Kroeze, J. M. F. H. Achten, K. Maatje, and W. Rossing (1997). Results of a multivariate approach to automated oestrus and mastitis detection. *Livestock Production Sci.* 48:219-227.

- Dempster, A. P., M. R. Selwyn, and B. J. Weeks (1983). Combining historical and randomized controls possessing trend properties. *J. Amer. Statistical Assoc.* 78:221-227.
- Diakosavvas, D. (1995). How integrated are world beef markets? The case of Australian and U. S. beef markets. *Agric. Economics* 12:37-53.
- Dimopoulos, I., S. Lek, and J. Lauga (1996). Rainfall-runoff modelling by neural networks and Kalman filter. *Hydrological Sci. J.* 41:179-193.
- Dixon, B. J. and R. E. Howitt (1979). Continuous forest inventory using a linear filter. *Forest Sci.* 25:675-689.
- Dubach, A. C. and H. Markl (1992). Application of an extended Kalman filter method for monitoring high density cultivation of *Escherichia coli*. *J. Ferment. Bioeng.* 73:396-402.
- Elliot, R. J. and V. Krishna-Murthy (1997). Exact finite-dimensional filters for maximum likelihood parameter estimation of continuous time linear Gaussian systems. *SIAM J. Control Optim.* 35:1908-1923.
- Eppstein, M. J. and D. E. Dougherty (1996). Simultaneous estimation of transmissivity values and zonation. *Water Resources Res.* 32:3321-3336.
- Erjavec, E. and J. Turk (1997). Supply elasticities in Slovene agriculture. *Zbornik Biotehniške Fakultete Univerze v Ljubljani, Kmetijstvo* No. 70, 85-98.
- Eskridge, K. M. and R. F. Mumm (1992). Choosing plant cultivars based on the probability of outperforming a check. *Theoretical and Applied Genetics* 84:894-900.
- Fahrmeir, L. (1992). Posterior mode estimation by extended Kalman filter for multivariate dynamic general models. *J. Amer. Statistical Assoc.* 87:501-509.
- Fahrmeir, L. and S. Wagensfal (1996). Smoothing hazard functions and time varying effects in discrete duration and competing risk models. *J. Amer. Statistical Assoc.* 91:1584-1594.
- Federer, W. T. (1951). Evaluation of variance components from a group of experiments with multiple classifications. *Res. Bull.* 380, *Agric. Expt. Sta., Iowa State College*.
- Federer, W. T. (1998) Recovery of interblock, intergradient, and intervariety information in incomplete block and lattice rectangle designed experiments. *Biometrics* 54:471-481.
- Federer, W. T., J. Crossa, and J. Franco (1998). New forms of spatial analyses with mixed model effects and exploratory model selection. BU-1406-M in the Technical Report Series of the Department of Biometrics, Cornell University, Ithaca, NY.
- Federer, W. T. and B. R. Murty (1987). Uses, limitations, and requirements of multivariate analysis for intercropping experiments. In *Biostatistics* (Ed: J. B. MacNeill and G. J. Umphrey), D. Reidel Publishing Co., Dordrecht, pp. 269-283.
- Finlay, K. W. and G. W. Wilkinson (1963). The analysis of adaptation in a plant breeding programme. *Australian J. Agric. Res.* 14:742-756.
- Fisher, R. A. (1932). Inverse probability and the use of likelihood. *Proc. Cambridge Phil. Soc.* 28:257-261.
- Flather, M. D. (1997). Strengths and limitations of meta-analysis: Larger studies may be more reliable. *Controlled Clinical Trials* 18:568-579.
- Franke, J. and A. Roeder (Editors) (1992). *Mathematical Modelling of Forest Ecosystems*. Proc. 5/27-31/91 Workshop, 174 pp.
- Freeman, S. N. and G. P. Kirkwood (1995). On a structural time series method for estimating stock biomass and recruitment from catch and effort data. *Fisheries Res.* (Amsterdam) 22:77-98.

- Frensham, A., B. Cullis, and A. Verbyla (1997). Genotype by environment variance heterogeneity in a two-stage analysis. *Biometrics* 53:1373-1383.
- Gauch, H. G. (1988). Model selection and validation for yield trials with interaction. *Biometrics* 44:705-715.
- Gave and Houston (1996). Amer. beech tree
- Gaver, D., D. Draper, P. Goel, J. Greenhouse, L. Hedges, C. Morris, and C. Waternaux (1992). Combining information: Statistical issues and opportunities for research. Panel Report, National Research Council, Washington, D. C., 217 pp.
- Geer, P. C. and P. V. Kloet (1983). Estimation of parameters in ground water flow problems using Kalman filter algorithms. In *Versl Maded Comm. Hydrol Onderz. TNO*, The Hague, The Commissie, May 1983, Issue 31, pp. 449-462.
- Geer, P. C. and P. V. Kloet (1985). Two algorithms for parameter estimation in ground water flow problems. *J. Hydrology* 77:361-378.
- Georgakakos, K. P. (1986) A generalized stochastic hydrometeorological model for flood and flash-flood forecasting. I. Formulation *Water Resources Res.* 22:2083-2095.
- Ghoul, M., M. Dardenne, C. Fontiex, and C. Marc (1991). An extended Kalman filtering technique for the on-line control of okt3 hybridoma cultures. *Biotechnol. Tech.* 5:367-370.
- Girah and C. D. Tankersley (1983).
- Givens, G. H., D. D. Smith, and R. L. Tweedie (1997). Publication bias in meta-analysis: A Bayesian data-augmentation approach to account for issues exemplified in the passive smoothing debate. *Statistical Science* 12:221-250.
- Gnanadesikan, R. and J. R. Kettenring (1988). A pragmatic view of multivariate methods applications. In *Statistics: An Appraisal* (ed: H. A. David and H. T. David). Proc. 50th Anniverary Conference , Iowa State University Press, Ames, Iowa, pp. 309-337.
- Gnanadesikan, R., J. R. Kettenring, and J. M. Landwehr (1982). Projection plots for displaying clusters. In *Statistics and Probability - Essays in Honor of C. R. Rao* (ed: G. Kallianpur, P. R. Krishnaiah, and J. K. Ghosh), pp. 269-280.
- Godfrey, J. T. and G. D. Foster (1996). Kalman filter method for estimating organic contaminant concentrations in major Chesapeake Bay tributaries. *Environmental Sci. Technology* 30:2312-2317.
- Goodall, E. A. and D. A. Sprevak (1985). A Bayesian estimation of the lactation curve of a dairy cow. *Animal Production* 40:189-193.
- Gordon, K. (1986). The multi-stage Kalman filter in medical monitoring. *Comput. Method. Program. Biomed.* 23:147-154.
- Gore, S. M. and B. A. Bradley (1988). *Developments in Nephrology* 21. Renal Transplantation Sense and Sensitization. Kluwer Academic Pub., Dordrecht, Boston, xxii + 311 pp.
- Gove, J. H. and D. R. Houston (1996). Monitoring the growth of American beech bark disease in Maine using Kalman filter. *Environmental Ecological Statistics* 3:167-187.
- Graham, W. and D. McLaughlin (1989). Stochastic analysis of nonstationary subsurface solute transport-3. Conditional moments. *Water Resources Res.* 25:2331-2355.
- Graham, W. D. and C. D. Tankersley (1993). Forecasting piezometric head levels in the Floridan aquifer: a Kalman filtering approach. *Water Resources Res.* 29:3791-3800.

- Gregoire, T. G. and D. K. Walters (1988). Composite vector estimates derived by weighting inversely proportional to variance. *Canadian J. Forest Res.* 18:282-284.
- Grewal, M. S. and A. P. Andrews (1993). *Kalman Filtering: Theory and Practice*. Prentice Hall, Englewood Cliffs, New Jersey, 381 pp.
- Grillenzoni, C. (1994). Optimal recursive estimation of dynamic models. *J. Amer. Statistical Assoc.* 89:777-787.
- Groton, E., C. Eagar, E. R. Cook, J. K. Ord, J. A. Derr, and R. A. J. Taylor (1988). Analyses of Great Smoky Mountain red spruce tree ring data. Gen Tech. Rep., Southern Forest Res. Expt. Sta., USDA Forest Service, No. SO-69, iii + 67 pp.
- Grum, M. (1998). Incorporating concepts from physical theory into stochastic modelling of urban runoff pollution. *Water Sci. Technology* 37:179-185.
- Gudmundsson G. (1995). Time series analysis of catch-at-length data. *ICES J. Marine Sci.* 52:781-795.
- Guiot, J. (1987). Late quaternary climatic change in France estimated from multivariate pollen time series. *Quat. Res.* 28:100-118.
- Gurewitch, J. and L. V. Hedges (1993). Meta-analysis: combining the results of independent experiments. In *Design and Analysis of Ecological Experiments* (Ed: S. M. Scheiner and J. Gurewitch), Chapman Hall, New York, pp 378-398.
- Hanks, R. J. (1992). Spatial and temporal soil water estimation considering soil variability and evapo transpiration uncertainty. *Water Resource Res.* 28:803-814.
- Harcum, J. B. and J. C. Loftis (1987). Spatial interpolation of Penman evapotranspiration. *Trans. ASAE* 30:129-136.
- Hayashi, Y., T. Shibasaki, R. Matsuda, and M. Uchiyama (1987). Real-time quantitative analysis of unresolved liquid chromatographic peaks with the use of the Kalman filter. *J. Chromatogr.* 407:59-64.
- Harcum, J. B. and J. C. Loftis (1987). Spatial interpolation of Penman evapotranspiration. *Trans. ASAE* 30:129-136.
- Hedges, L. V. (1992). Modeling publication selection effects in meta-analysis. *Statistical Science* 7:246-255.
- Higgins, J. P. T. (1996). Borrowing strength from extremal trials in meta-analysis. *Statistics in Medicine* 15:2733-2749.
- Hoo, K. A. and J. C. Kantor (1986). Linearization and control of a mixed-culture bioreactor with competition and external inhibition. *Math. Biosci.* 82:43-62.
- Ichihara, K. (1987). Analysis of run off from a logging road. *J. Jap. Forest Sci.* 69:329-342.
- Ikeda, Y. (1990). Quality evaluation of agricultural products by their dynamic properties. Description of dynamic properties by an autoregressive-moving average model and its application to quality evaluation. *J. Japanese Soc. Agric. Machinery* 52:43-49.
- Islam, M. N., A. Nagai, and A. Yomota (1994). Real-time flood forecasting in mountainous river basins with long- and short-term runoff model. *J. Irrigation Engineering Rural Planning* No. 26:31-47.
- Iyengar, S. and J. B. Greenhouse (1988 or 1992, Page 4). Selection models and the file drawer problem. *Statistical Science* 3:133-135.

- Jain, R. K. (1995). Comparative performance of state-space model based and ARIMA model-based methods of seasonal adjustment. *Proc. Amer. Statistical Assoc., Business and Economics Sec.*, pp 262-267.
- Jetto, L. (1985). Small-computer procedure for optimal filtering of hemodynamic data. *Biol. Eng. Comput.* 23:203-208.
- Jones, B. B. and W. F. Tompkins (1998). A physicist's guide to Kalman filters. World Wide Web, March, PDF(1.0/PS292k).
- Jones, P. N. and A. N. Pettitt (1992). Comparison of EEGs before and after stunning of cattle taking account of animal-to-animal variation. *Biometrical J.* 34:812-825.
- Jones, R. H. and L. M. Ackerson (1990). Serial correlation in unequally spaced longitudinal data. *Biometrika* 77:721-732.
- Jones, R. H. and F. Boadi-Boateng (1991). Unequally spaced longitudinal data with ARI serial correlation. *Biometrics* 47:161-176.
- Jones, R. M., J. F. Macgregor, and K. L. Murphy (1989). State estimation in wastewater engineering application to an anaerobic process. *Workshop On Statistical Methods for the Assessment of Point Source Pollution, Burlington, Ontario, 9/12-14/89*, pp.271-282.
- Kadoya, M. and H. Tanakamaru (1994). Real-time flood runoff forecasting with long- and short-term runoff model. ??
- Kalivianakis, M., S. L. J. Mous, and J. Grasman (1994). Reconstruction of the seasonally varying contact rate for measles. *Mathematical Biosciences* 124:225-234.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *J. Basic Engineering* 82:34-45.
- Kalman, R. E. and R. S. Bucy (1961). New results in linear filtering and prediction theory. *J. Basic Engineering* 83:95-108.
- Kannan, D. and Q. Zhang (1998). Non-linear filtering of an iterative multiplemodel with small observation noise: numerical methods. *Stochastic Analysis and Applications* 16:631-65?.
- Kaylen, M. S. and S. S. Koroma (1991). Trend, weather variables, and the distribution of U.S. corn yields. *Rev. Agric. Economics* 13:249-258.
- Kimura, D. K., J. W. Balsiger, and D. H. Ito (1996). Kalman filtering of delay-difference equation: Practical approaches and simulations. *Fishery Bulletin (Washington, D. C.)* 94:678-691.
- Kitanidis, P. K. and R. L. Bras (1980). Adaptive filtering through detection of isolated transient errors in rainfall runoff models *Water Resources Res.* 16:740-748.
- Knapp, K. C. and K. Konyar (1991). Perennial crop supply response: a Kalman filter approach. *Am. J. Agric. Econ.* 79:841-849.
- Kohn, R. and C. F. Ansley (1987). Comment on Kitagawa article. *J. Amer. Statistical Assoc.* 82:1041-1044.
- Larose, D. T. (1997). Grouped random effects models for Bayesian meta-analysis. *Statistics in Medicine* 16:1817-1829.
- Leblanc, D. C. (1993). Temporal and spatial variation of oak growth-climate relationships along a pollution gradient in the midwestern United States. *Canadian J. Firest Res.* 23:772-782.
- Lee, S. C., Y. B. Hwang, H. N. Chang, and Y. K. Chang (1991). Adaptive control of dissolved oxygen concentration in a bioreactor. *Biotechnol. Bioeng.* 37:597-607.

- Leondes, C. T. (1982). Advances in the techniques and technology of the application of non-linear filters and Kalman filters. NATO Advisory Group for Aerospace Research, Neuilly-swi-Seine, France, No. 256.
- Letotsky, S., Mikulevicius, R. and B. L. Rozovskii (1997). Non-linear filtering revisited: A spectral approach. *SIAM J. Control Optim.* 35:435-461.
- Lewis, W. H. Jr. and S. C. Rutan (1991). Guanidinium-induced differential kinetic denaturation of alkaline phosphatase isozymes. *Anal. Chem.* 63:627-629.
- Li, Shi, and Roth (1994). *Communications in Statistics*
- Lin, J. S. (1994). Extraction of dynamic soil properties using extended Kalman filter. *J. Geotechnical Eng.* 120:2100-2117.
- Lin, X., X. Zhang, and X. Xin (1996). Application of Kalman filter in agricultural economics forecasting. *Proc. AEEE Int'l Conf. on Systems, Ma??, and Cybernetics, Beijing*, 3:2185-2189.
- Liu, Y., D. Zhao, R. Altman, and O. Jardetsky (1992). A systematic comparison of three structure determination methods from nmr data dependence upon quality and quantity of data. *J. Biomol. Nmr* 2:373-388.
- Lototsky, S., R. Mikulevicius, and B. L. Rozowskic (1997). Non-linear filtering revisited: A special approach. *SIAM J. Control Optim.* 35:435-461.
- MacLachlan, G. J. and K. E. Basford (1988). *Mixture Models: Inferences, and Application of Clustering*. Marcel Dekker, Inc., New York, 253 pp.
- Makinen, H. (1997). Reducing the effect of disturbance on tree-ring data using intervention detection. *Scandinavian J. Forest Res.* 12:351-355.
- Malaterre, P. and O. Pilote (1998). Linear quadratic optimal controller for irrigation canals. *J. Irrigation Drainage Eng.* 124(4):187-194.
- Mandenius, C. F., B. Mattiasson, J. P. Axelsson, and P. Hagander (1987). Control of an ethanol fermentation carried out with alginate entrappe *saccharomyces-cerevisine*. *Biotechnol. Bioeng.* 29:941-949.
- Marchant, J. A., T. Hague, and N. D. Tillett (1997). Row-following accuracy of an autonomous vision-guided agricultural vehicle. *Computers Electronics Agric.* 16:165-175.
- McLachlan, G. J. and K. E. Basford (1988). *Mixture Models - Inference and Applications to Clustering*. Marcel Dekker, Inc. New York and Basel.
- Meinhold, R. J. and N. D. Singpurwalla (1983) Understanding Kalman filter. *Amer. Statistician* 37:123-127.
- Meinhold, R. J. and N. D. Singpurwalla (1989). Robustification of Kalman filter models. *J. Amer. Statistical Assoc.* 84:727-732.
- Mendelssohn, R. (1988). Some problems in estimating population sizes from catch-at-age data. *U.S. Natl. Mar. Fish Serv. Fish Bull.* 86(4):617-630.
- Mimikou, M. A., P. S. Hatjisava, Y. S. Kouvopoulos, and E. N. Anagnostou (1992). The influence of basin aridity on the efficiency of runoff predicting models. *Nordic Hydrology* 23:105-120.
- Mocan, H. N. and K. Topyan (1993). Illicit drug use and health: analysis and projections of New York City birth outcomes using a Kalman filter model. Working paper, National Bureau of Economic Research, Cambridge, MA, No. 4359.

- Montana, D. J. (1995). Strongly typed genetic programming. *Evolutionary Computation* 3:199-230.
- Mosteller, F. and T. C. Chalmers (1992). Some progress and problems in meta-analysis clinical trials. *Statistical Science* 7:227-236.
- Mulholland, M. (1989). An autoregressive atmospheric dispersion model for fitting combined source and receptor data sets. *Atmos. Environ.* 23:1443-1458.
- Murase, H., Y. Nishiura, and K. Mitani (1997). Environmental control strategies based on plant responses using intelligent machine vision technique. *Computers Electronics Agric.* 18:137-148.
- Murase, H., R. Yamauchi, and N. Honami (1992). Simultaneous control of temperature and humidity in a confined space using Kalman filter neural networks system identification. *Environ. Control Biol.* 30:37-44.
- Nahlik, J. and Z. Burianec (1988). On-line parameter and state estimation of continuous cultivation by extended Kalman filter. *Appl. Microbiol. Biotechnol.* 28:128-134.
- Normand, S. and D. Trichler (1992). Parameter updating in a Bayes network. *J. Amer. Statistical Assoc.* 87:1109-1115.
- Ohtuska, K. (1998). Better M-sequence design for particular identification based on extended Kalman filter. *IEE Proc. Control Theory and Applications* 145:92-96.
- Olkin, I. (1992). Meta-analysis: Methods for combining independent studies. *Statistical Science* 7:226.
- Olkin, I. and A. Sampson (1998). Comparison of meta-analysis versus analysis of individual patient data. *Biometrics* 54:317-323.
- Oltgen, J. W. and F. N. Owens (1987). Beef cattle feed intake and growth: empirical Bayes derivation of the Kalman filter applied to a nonlinear dynamic model. *J. Animal Sci.* 65:1362-1370.
- Pan, Y. and D. J. Raynal (1995). Predicting growth of plantation conifers in the Adirondack Mountains in response to climate change. *Canadian J. Forest Res.* 25:48-56.
- Pan, Y., S. Xia, S. Si, M. Zhang, L. Shi, and X. Liu (1990). Target factor analysis UV spectrophotometry for the simultaneous determination of six amino acids. *J. Chemometrics* 4:323-330.
- Papakyriazis, A. and P. Papakyriazis (1998). Optimal environmental policy under imperfect information. *Kybernetika* 27:137-154.
- Papanastassiou, D. (1992). A generalized state-space model and its application to a set of extended Kalman filter equations. *Communications in Statistics (Theory and Methods)* 21:3255-3262.
- Patsi, J., L. Malkki, and S. Tammilehto (1992). Simultaneous determination of caffeine propyphenazone and salicylamide by Kalman filter spectrophotometry and high-performance liquid chromatography. *Acta Pharm. Nord.* 4:69-72.
- Pella, J. J. (1993). Utility of structural time series models and the Kalman filter for predicting consequences of fishery actions. *Alaska Sea Grant College Report*, No. 93(2).
- Pena, D. (1997). Combining information in statistical modeling. *American Statistician* 51:326-332.
- Penczek, P., W. Grochulski, J. Grzyb, and M. Kowalczyk (1987). The use of a multi-channel Kalman filter algorithm in structural analysis of the epileptic EEG. *Int. J. Bio-Med. Comput.* 20:135-151.

- Penczek, P., W. Grochulski, and M. Kowalczyk (1987). Computer-aided analysis of the epileptic EEG. *Acta Physiol. Pol.* 37:262-274.
- Pentland, A. P. (1992). Dynamic vision. In *Neural Networks for Vision and Image Processing* (eds: G. A. Carpenter and S. Grossberg). MIT Press, Cambridge, MA, pp. 133-159.
- Pfeffermann, D. (1991). Estimation and seasonal adjustment of population means using data from repeated surveys. *J. Economic Business Stat.* 9:163-175.
- Pfeffermann, D., M. Feder, and D. Signorelli (1998). Estimation of autocorrelations of survey errors with application to trend estimation in small areas. *J. Business and Economic Statistics* 16:339-348.
- Piegorsch, W. W., E. P. Smith, D. Edwards, and R. L. Smith (1998). Statistical advances in environmental science. *Statistical Science* 13:186-208.
- Piepho, H. P. (1995). Detecting and handling heteroscedasticity in yield trial data. *Communications in Statistics (Simulation)* 21:243-274.
- Piepho, H. P. (1997). Analyzing genotype-environment data by mixed models with multiplicative terms. *Biometrics* 53:761-766.
- Piepho, H. P. (1996). A simplified procedure for comparing the stability of cropping systems. *Biometrics* 52:315-320.
- Poole, R. W. (1978). The statistical prediction of population fluctuations. *Annual Rev. Ecology Systematics* 9:427-448.
- Rao, C. R. (1982). Diversity and dissimilarity coefficients: A unified approach. *Theor. Population Biology* 21:24-42.
- Rao, M. J. M. (1987). *Filtering and Control of Macro-Economic Systems*. North Holland/Elsevier Sci. Publications, 275 pp.
- Rao, P. and H. K. Shimizu (1992). Meta-analysis and trend estimation of acute myocardial infarction data. *Statistics in Medicine* 10:149-162.
- Read, W. J. and C. M. Simons (1996). Analyzing catch-effort data by means of the Kalman filter. *Canadian J. Fish Aquatic Sci.* 53:2157-2166.
- Reddy, J. M. (1994). Stochastic optimal control of irrigation channels. In *Advances in Water Resources Technology and Management, Proc. Second European Conf., Lisbon, 6/14-18/94*, pp.189-195.
- Romera, R. and T. Cipra (1995). On practical implementation of robust Kalman filtering. *Communications in Statistics (Simulation)* 24:461-488.
- Rotunno, T. (1992). Combined Kalman filtering and steepest descent minimization for quantitative analysis of unresolved misaligned chromatographic peaks. The resolution of alternariol and altenuisol mycotoxins in mixtures by HPLC. *Chromatographia* 34:56-62.
- Roush, W. B., K. Tomiyama, K. H. Garnabui, T. H. D. Alfonso, and T. L. Cravener (1992). Kalman filter and an example of its use to detect changes in poultry production responses. *Computers Electronics Agric.* 6:347-356.
- Rutan, S. C. and C. B. Motley (1987). Factor analysis and Kalman filter studies of severely overlapped amino acid derivatives in thin layer chromatography. *Anal. Chem.* 59:2045-2050.
- Sauter, D., B. J. Martin, N. Di Renzo, and C. Vomscheid (1991). Analysis of eye tracking movements using innovations generated by a Kalman filter. *Med. Biol. Eng. Comput.* 29:63-69.

- Schlain, B. R., P. T. Lavin, and C. L. Hayden (1992). Using an autoregressive model to detect departures from steady states in unequally spaced tumour biomarker data. *Stat. Med.* 11:515-532.
- Schlink, U., O. Herbarth, and G. Tetzlaff (1997). A component time-series model for SO₂ data: Forecasting, interpretation and modification. *Atmospheric Environment* 31:1285-1295.
- Schneider, W. (1987). The role of Kalman filters as an aid to statistical analysis in econometric models. *Arbeiten Institut Statistik Okonometrie Christian-Albrechts Univ. Kiel*, No. 36, 31 pp.
- Schnute, J. T. (1994). A general framework for developing sequential fisheries models. *Canadian J. Fisheries Aquatic Sci.* 51:1676-1688.
- Schmalter, S. F. (1994). Applied state space modelling of non-Gaussian time series using integration-based Kalman filtering. *Statistics and Computing* 4:259-269.
- Seldon, B. J. and R. Boyd (1991). An examination of the stability of regional lumber demand in the contiguous United States, 1950-1985. *Annals Regional Sci.* 25:209-215.
- Sen, Z. (1984). Adaptive pumping test analysis. *J. Hydrology* 74:259-270.
- Shang, H. W. and C. M. Liago (1996). A Kalman filter applied to alarm dissolved oxygen concentration flux in fish pond. *J. Fisher. Soc., Tiawan* 23:17-29.
- Shi, L., Z. Li, Z. Xu, Z. Pan, and L. Wang (1991). Simultaneous analysis of cobalt-II nickel-II copper-II zinc-II and cadmium-II by spectrophotometry and the Kalman filter. *J. Chemometrics* 5:193-200.
- Shi, L., Z. Xu, Z. Pan, X. Liu, Y. Gao, F. Qian, S. Nie, and Z. Li (1990) Application of the Kalman filter to the simultaneous determination of four amino acids by direct UV spectrophotometry. *J. Micronutr. Anal.* 8:1-12.
- Shi, Y., D. D. Y. Ryu, and W. K. Yuan (1993). Effect of oxygen and ethanol on recombinant yeast fermentation for hepatitis B virus surface antigen production. *Biotech. Bioengineering* 41:55-66.
- Shiina, T., K. Ikeda, and M. Saito (1987). Estimation of tissue parameters derived from reflected ultrasound. *Med. Prog. Technol.* 12:185-196.
- Shimitzu, K. (1993). An overview of the control system design of bioreactors. *Advances in Biochemical Engineering/Biological Chemistry. Biotechnology* 50:65-84.
- Shimitzu, H., K. Araki, S. Shioya, and K. Suga (1991). Optimal production of glutathione by controlling the specific growth of yeast in fed-batch culture. *Biotechnol. Bioeng.* 38:196-205.
- Shimitzu, H., S. Shioya, K. I. Suga, and T. Takamatsu (1989). Profile control of the specific growth rate in fed-batch experiments. *Appl. Microbiol. Biotechnol.* 30:276-282.
- Shimitzu, H., T. Takamatsu, S. Shioya, and K. I. Suga (1989). An algorithmic approach to constructing the on-line estimation system for the specific growth rate. *Biotechnol. Bioeng.* 33:354-364.
- Shukla, G. K. (1972). Some statistical aspects of partitioning genotype-environment components of variability. *Heredity* 29:237-245.
- Simutis, R., I. Havelik, and A. Luebbert (1992). A fuzzy-supported extended Kalman filter: A new approach to state estimation and prediction exemplified by alcohol fermentation in beer brewing. *J. Biotechnology* 24:211-234.

- Sittig, D. F. and K. H. Cheung (1992). A parallel implementatin of a multi-state Kalman filtering algorithm to detect EEG arrhythmias. *Int. J. Clin. Monit. Comput.* 9:1-10.
- Sittig, D. F. and M. Factor (1990). Physiologic trend detection and artifact rejection. A parallel implementation of multi-state Kalman filtering algorithm. *Comput. Methods Program. Biomed.* 31:1-10.
- Skagen, D. W. (1988). Estimation of running frequency spectra using a Kalman filter algorithm. *J. Biomed. Eng.* 10:275-279.
- Sloan, W. T., A. Jenkins, and A. Eatherall (1994). A simple model of stream nitrate concentrations in forested and deforested catchments in Mid-Wales. *J. Hydrology (Amsterdam)* 158:61-78.
- Slordahl, S. A., J. E. Solbakken, H. Piene, B. A. J. Angelsen, O. Rossvoll, and S. O. Samstad (1990). Quantification of aortic regurgitation by doppler echocardiography. A new method evaluated in pigs. *Med. Biol. Eng. Comput.* 28:300-305.
- Smith, A. F. M. and M. West (1983). Monitoring renool?? transplants: An application of multi-process Kalman filter. *Biometrics* 39:867-878.
- Spall, J. C. (1988). *Bayesian Analysis of Time Series and Dynamic Models*. Marcel Dekker, Inc., New York, 536 pp.
- Sprague, G. F. and W. T. Federer (1951). A comparison of variance components in corn yield trials: II. Error, year variety, location variety, and variety components. *Agronomy J.* 43:535-541.
- Storm, B., K. H. Jenson, and J. C. Refsgaard (1988). Estimation of catchment rainfall uncertainty and its influence on runoff prediction. *Nordic Hydrology* 19:77-88.
- Studer, R. M., R. J. P. de Figueiredo, and G. S. Moschytz (1984). An algorithm for sequential signal estimation and system identification for electro myogram signals. *Trans. Biomed. Eng.* 31:285-295.
- Sullivan, P. J. (1992). A Kalman filter approach to catch-at-length analysis. *Biometrics* 48:237-257.
- Sun, X. D., Y. Wang, Z. M. Xu, and C. S. Xu (1996). Research spatial non-homogeneous Kalman filter and its applications in cephalometric image. *Proc. 3rd Int'l Conf. Signal Processing*, 2:1078-1081.
- Surendran, S. and R. J. Mulholland (1986). Estimation of atmospheric carbon dioxide concentration using Kalman Filtering. *Int. J. Syst. Sci.* 17:897-910.
- Svrcek, W. Y., R. F. Elliot, and J. E. Zajic (1974). The extended Kalman filter applied to a continuous culture model. *Biotechnol. Bioeng.* 16:827-846.
- Ta-Hsin Li and B. Kedem (1998). Tracking abrupt frequency changes. *J. Time Series Analysis* 19:69-82.
- Tadros, M. T. Y. and M. A. M. Shaltout (1989). Effect of Hanning and Parzen windows and Kalman filter on the spectral analysis of solar constant and solar activities. *Solar Energy* 43:331-338.
- Talhami, H. E. and R. I. Kitney (1988). Maximum likelihood frequency tracking of the audio pulsed doppler ultrasound signal using a Kalman filter. *Ultrasound Med. Biol.* 14:599-610.
- Tan, W. and Z. Xiang (1998). A state-space model for the HIV epidemic in homosexual populations and some applications. *Mathematical Biosciences* 152:29-61.

- Tani, A., H. Murase, M. Kiyota, and N. Honami (1992). Growth simulation of alfalfa cuttings in vitro by Kalman filter neural network. *Acta Horticulturae* 2:671-676.
- Tanizaki, H. and R. S. Mariano (1994). Prediction, filtering and smoothing in non-linear and non-normal cases using Monte Carlo integration. *J. Applied Economics* 9:163-180.
- Taplin, R. H. (1998). The reverse Kalman filter. *Communications In Statistics, Theory and Methods* 27:2547-2558.
- Thompson, S. G. (1997). Investigating underlying risk as a source of heterogeneity in meta-analysis. *Statistics in Medicine* 16:2741-2758.
- Thysen, I. (1993). Monitoring bulk tank cell counts by a multi-process Kalman filter. *Acta Agric. Scandinavica* 43(A):58-64.
- Tiwari, R. C. (1994). A non-parametric Bayesian approach to analysis of drug dependence data. *Communications in Statistics (Theory and Methods)* 23:625-642.
- Tiwari, R. C., Y. Yang, and J. N. Zalkikar (1996). Time series analysis of BOD data using the Gibbs sampler. *Environmetrics* 7:567-578.
- Trimble, I. M., M. West, M. S. Knapp, R. Pownall, and A. F. M. Smith (1983). Detection of renal allo graft rejection by computer. *Ed. J.* 286:1695-1699.
- Tucci, M. P. (1997). Adaptive control in the presence of time varying parameters. *J. Economy and Dynamic Control* 22:39-47.
- Turk, J., E. Erjavec, and D. Gambelli (1997). Kalman filter application in estimating agricultural supply. *Proc. 4th Int'l Symposium Operational Res. in Slovene, Preddvor, 10/1-3/97*, pp. 239-243.
- Van Bebber, J., N. Reinsch, W. Junge, and E. Kalm (1997). monitoring dairy milk yield records with a recursive model (Kalman filter). *Proc. 48th Annual Mtg. European Assn. Animal Prod., Vienna, August 1997*, 283 pp.
- Van den Brakel, J. A. and H. Visser (1996). The influence of environmental conditions on tree-ring series of Norway spruce for different canopy and vitality classes. *Forest Sci.* 42:206-219.
- Van den Bergh, J. C. J. H., . J. Bulton, P. Nijkamp, and P. Peppring (1997). *Meta-Analysis Versus Analysis of Variance of Individual Patient Data*. Kluwar Academic Publications, Dordrecht/Boston, 219 pp.
- Van Deusen, P. C. (1987). Pollution and tree rings.
- Van Deusen, P. C. (1989). A model-based approach to tree ring analysis. *Biometrics* 45:763-780.
- Van Deusen, P. C. (1990). Evaluating time-dependent tree ring and climate relationships. *J. Environ. Quality* 19:481-488.
- Van Deusen, P. C. (1991). Population formulations for modeling tree rings from climate: A unified approach. *J. Environ. Qual.* 20:823-827.
- Van Eeuwijk, F. A. (1995). Multiplicative interaction in generalized linear models. *Biometrics* 51: 1017-1032.
- Van Eeuwijk, F. A. (1992). Interpreting genotype by environment interaction using redundancy analysis. *Theoretical and Applied Genetics* 85:89-100.
- Vanhonacker, W. R. (1996). Meta-analysis and response surface extrapolation: A least squares approach . *American Statistician* 50:294-297.
- Verducci, J. S. (1988). Combining information from similar experiments. *Biometrical J.* 30:387-395.

- Visser, H. (1986). Analysis of tree ring data using the Kalman filter technique. Symp. Effects Environmental Pollution on Wood Structure Quality, 18th IUFRO World Congress, Ljubljana, 9/10-11/86, pp. 289-298.
- Visser, H. and J. Molenaar (1990). Estimating trends in tree-ring data. *Forest Sci.* 36:87-100.
- Visser, H., F. Noppert, H. Van Wakeren, And J. Vaessen (1989). Xylem sap velocity in relation to weather and air pollution. *Intl. Assoc. Wood Anat. Bull.* 10:427-439.
- Viswanathan, M. N. and D. A. Evans (1983). Estimation of groundwater recharge using Kalman filter method. Hunter District Water Board, New South Wales, pp. 200-203.
- Wachler, K. N. and M. L. Straf (1990). *The Future of Meta-Analysis*. Russell Sage Foundation, New York.
- Walters, D. K., H. E. Burkhart, M. R. Reynolds, Jr., and T. G. Gregoire (1991). A Kalman filter approach to localizing height-age equations. *Forest Sci.* 37:1526-1537.
- Wang, Y. and B. Payandeh (1994). Application of the Kalman filter model in site index equation construction. *Canadian J. Forest Res.* 24:1415-1418.
- Wells, C. (1996). *The Kalman Filter in Finance*. Kluwer Academic Publ., London, 169 pp.
- Wendroth, O., G. G. Katul, M. B. Parlange, C. E. Puente, and D. R. Nielsen (1993). A nonlinear filtering approach for determining hydraulic conductivity functions in field soils. *Soil Sci.* 156:293-301.
- Whitehead, P., B. Beck, and E. A. O'Connell (1981). A systems model of stream flow and water quality in the Bedford Ouse River UK system 2. Water quality modeling. *Water Res.* 15:1157-1172.
- Whittaker, J. and S. Fruhwirth-Schatter (1993). A dynamic point change model for detecting the onset of growth in bacteriological infections. *Applied Statistics* 43:625-640.
- Wild, P. and A. Kozakieurecz (1993). Kalman filter method in the analysis of vibration due to water waves. *World Scientific, London*, 1-173.
- Wilson, P. H. and B. E. McKenzie (1998). Information processing deficit associated with developmental coordination disorder (DCD): A meta-analysis of research findings. *J. Child Psychology Psychiat.* 39:829-840.
- Woolfson, M. S. (1991). Study of cardiac arrhythmia using the Kalman filter. *Med. Biol. Eng. Comput.* 29:398-405.
- Wu, W., S. C. Rutan, A. Baldovin, and D. L. Massart (1996). Feature selection using the Kalman filter for classification of multivariate data. *Analytica Chimica Acta* 335:11-22.
- Wu, X. and K. H. Bellgardt (1998). Fast on-line evaluation of flow injection analysis signals based on parameter estimation by an extended Kalman filter. *J. Biotechnology* 62:11-28.
- Xiong, R., A. Velasco, M. Silva, and B. D. Perez (1991). Performance of the Kalman filter algorithm in differential reaction-rate methods. *Anal. Chim. Acta* 251:313-320.
- Yada, T. and T. Kitagawa (1995). Analysis for abrupt changes of trend in time series under business competition. *Proc. Amer. Statistical Assoc., Business and Economics Sec.*, pp 66-71.
- Yasuda, H., K. Takuma, and K. Sekikawa (1994). Runoff characteristics of small mountainous basins. *Bulletin Faculty Agric., Tottori Univ.*, 47-43-49

- Yates, F. and W. G. Cochran (1938). The analysis of groups of experiments. *J. Agric. Science* 28:556-580.
- Yu, Y. S., M. Heidari, and G. T. Wang (1989). Optimal estimation of contaminant transport in groundwater. *Water Resources Bulletin* 25:295-300.
- Yude, P. and R. J. Dudley (1995). Predicting growth of plantation conifers in Adirondack Mountains in response to climate change. *J. Forest Res.* 25:48-56.
- Zavaleta, L. R. and B. L. Dixon (1982). Economic benefits of Kalman filtering for insect pest management. *J. Economic Entomology* 75:982-988.